Helsinki Graduate Program in Biotechnology and Molecular Biology (GPBM)/
Integrative Life Sciences (ILS) Doctoral Program

THE ROLE OF *MED12* IN TUMORIGENESIS

Kati Kämpjärvi

Department of Medical and Clinical Genetics, Medicum
&
Genome-Scale Biology Research Program, Research Programs Unit
Faculty of Medicine
University of Helsinki
Helsinki, Finland

Academic dissertation

To be publicly discussed, with the permission of the Faculty of Medicine, University of Helsinki, in Haartman Institute, Lecture Hall 1, Haartmaninkatu 3, Helsinki, on the 18th of November 2016, at 12 noon.

Helsinki 2016

Supervised by

Docent Pia Vahteristo, Ph.D.
Department of Medical and Clinical Genetics, Medicum
Genome-Scale Biology Research Program, Research Programs Unit
Faculty of Medicine
University of Helsinki
Helsinki, Finland

Academy Professor Lauri A. Aaltonen, M.D., Ph.D.
Department of Medical and Clinical Genetics, Medicum
Genome-Scale Biology Research Program, Research Programs Unit
Faculty of Medicine
University of Helsinki
Helsinki, Finland

Reviewed by

Docent Miina Ollikainen, Ph.D.
Institute for Molecular Medicine Finland, FIMM
Department of Public Health, Clinicum
Faculty of Medicine
University of Helsinki
Helsinki, Finland

Docent Katri Pylkäs, Ph.D.
Laboratory of Cancer Genetics and Tumor Biology
Biocenter Oulu
Faculty of Medicine
University of Oulu
Oulu, Finland

Official opponent

Professor Anne Kallioniemi, M.D., Ph.D.
BioMediTech
University of Tampere
Tampere, Finland

ISBN 978-951-51-2643-6 (paperback)
ISBN 978-951-51-2644-3 (PDF)
htpp://ethesis.helsinki.fi
Unigrafia
Helsinki 2016

TABLE OF CONTENTS

ORIGINAL PUBLICATIONS

This thesis is based on the following original publications which are referred to in the text by their roman numerals I-V:

I **Kämpjärvi K***, Mäkinen N*, Kilpivaara O, Arola J, Heinonen H-R, Böhm J, Abdel-Wahab O, Lehtonen HJ, Pelttari LM, Mehine M, Schrewe H, Nevanlinna H, Levine RL, Hokland P, Böhling T, Mecklin JP, Bützow R, Aaltonen LA, Vahteristo P. Somatic *MED12* mutations in uterine leiomyosarcoma and colorectal cancer. *British Journal of Cancer* 2012, 107:1761-1765.

II **Kämpjärvi K**, Park MJ, Mehine M, Kim NH, Clark AD, Bützow R, Böhling T, Böhm J, Mecklin JP, Järvinen H, Tomlinson IPM, van der Spuy ZM, Sjöberg J, Boyer TG, Vahteristo P. Mutations in exon 1 highlight the role of *MED12* in uterine leiomyomas. *Human Mutation* 2014, 35:1136-1141.

III **Kämpjärvi K**, Mäkinen N, Mehine M, Välipakka S, Uimari O, Pitkänen E, Heinonen HR, Heikkinen T, Tolvanen J, Ahtikoski A, Frizzell N, Sarvilinna N, Sjöberg J, Bützow R, Aaltonen LA, and Vahteristo P. *MED12* mutations and *FH* inactivation are mutually exclusive in uterine leiomyomas. *British Journal of Cancer* 2016, 114:1405-1411.

IV **Kämpjärvi K***, Järvinen TM*, Heikkinen T, Ruppert AS, Senter L, Hoag KW, Dufva O, Kontro M, Rassenti L, Hertlein E, Kipps TJ, Porkka K, Byrd JC, de la Chapelle A, Vahteristo P. Somatic *MED12* mutations are associated with poor prognosis markers in chronic lymphocytic leukemia. *Oncotarget* 2015, 6:1884-1888.

V Heikkinen T*, **Kämpjärvi K***, Keskitalo S*, von Nandelstadh P, Liu X, Rantanen V, Pitkänen E, Kinnunen M, Kuusanmäki H, Kontro M, Turunen M, Mäkinen N, Taipale J, Heckman C, Lehti K, Mustjoki S, Varjosalo M, Vahteristo P. Somatic *MED12* nonsense mutation escapes mRNA decay and reveals a motif required for nuclear entry. *Submitted*.

*Equal contribution

ABBREVIATIONS

2SC	S-(2-succinyl)-cysteine
A	adenine/alanine
aa	amino acid
ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
AP	affinity purification
C	cytosine/cysteine
C-terminus	carboxy-terminus
CCNC	cyclin C
CDK8	cyclin-dependent kinase 8
CDK19	cyclin-dependent kinase 19
cDNA	complementary deoxyribonucleic acid
CLL	chronic lymphocytic leukemia
COL4A5	collagen type IV alpha 5
COL4A6	collagen type IV alpha 6
COSMIC	Catalogue of Somatic Mutations in Cancer
CPG	cancer predisposing gene
CTD	carboxy-terminal domain
CUX1	cut-like homeobox 1
D	aspartic acid
del	deletion
DNA	deoxyribonucleic acid
E	glutamic acid
ECM	extracellular matrix
FF	fresh frozen
FFPE	formalin-fixed paraffin-embedded
FH	fumarate hydratase
G	guanine/glycine
Gli3	glioma-associated oncogene family zinc finger 3
H	histidine/histone
HE	hematoxylin-eosin
HIF1A	hypoxia-inducible factor 1 alpha
HLRCC	hereditary leiomyomatosis and renal cell cancer
HMGA1	high mobility group AT-hook 1
HMGA2	high mobility group AT-hook 2
HPV	human papillomavirus
ICGC	International Cancer Genome Consortium
IF	immunofluorescence
IGHV	immunoglobulin heavy chain variable
IHC	immunohistochemistry
ins	insertion
K	lysine
L	leucine
LC-MS	liquid chromatography-mass spectrometry
lncRNA	long non-coding RNA
LOH	loss of heterozygosity
M	methionine
MED12	mediator complex subunit 12
MED12L	mediator complex subunit 12-like
MED13	mediator complex subunit 13
MED13L	mediator complex subunit 13-like
MIM	Mendelian Inheritance in Man
miRNA	micro-ribonucleic acid
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MSI	microsatellite instability
N	asparagine
NLS	nuclear localization signal
NMD	nonsense mediated decay
NOTCH1	Notch (Drosophila) homolog 1, translocation-associated
NPC	nuclear pore complex
N-terminus	amino terminus
OPA	opposite paired domain
p	short arm of a chromosome
P	proline
PCP	planar cell polarity
PCR	polymerase chain reaction
Pol II	RNA polymerase II
q	long arm of a chromosome
Q	glutamine
R	arginine
RAD51B	RAD51 paralog B
REST	RE1-silencing transcription factor
RNA	ribonucleic acid
S	serine
Shh	Sonic hedgehog
STUMP	smooth muscle tumor with uncertain malignant potential
T	thymine/threonine
T-ALL	T-cell acute lymphoblastic leukemia
TCGA	The Cancer Genome Atlas
TERT	telomerase reverse transcriptase
TGF-β	transforming growth factor beta
TGF-βR2	transforming growth factor beta receptor 2
TMA	tissue microarray
TP53	tumor protein 53
TSG	tumor suppressor gene
V	valine
W	tryptophan
WB	Western blotting
WES	whole exome sequencing
WGS	whole genome sequencing
WHO	the World Health Organization
Wnt	wingless-related integration site
WT	wild type
X	X chromosome/stop
XLID	X-linked intellectual disability
ZAP-70	70 kD zeta-associated protein

Gene names and symbols are italicized in the text.

ABSTRACT

Mediator complex subunit 12 (MED12), encoded by the *mediator complex subunit 12* gene on Xq13.1, is a component of a conserved Mediator complex and an important player in the regulation of general and gene-specific transcription. Mediator is a multiprotein complex required for ribonucleic acid (RNA) polymerase II-dependent transcription of most protein coding genes. MED12 composes the kinase module of the Mediator together with Mediator complex subunit 13 (MED13), Cyclin C (CCNC) and Cyclin-dependent kinase 8 or 19 (CDK8/19). With variable association to the Mediator core, it regulates the activity of the complex and also participates in scaffold formation and transcription elongation. Somatic *MED12* mutations were implicated in human tumorigenesis for the first time when exome sequencing of uterine leiomyomas identified extremely frequent mutations in the gene. All the observed mutations were situated in an evolutionarily conserved area in exon 2 and the preceding intron-exon boundary. Mutations were missense changes affecting specific mutation hot spots and small in-frame insertions and deletions at the same genomic region. Uterine leiomyomas are benign smooth muscle tumors estimated to affect up to 77% of reproductive-age women. Despite their benign nature, clinically relevant lesions can cause a variety of symptoms, and, consequently, uterine leiomyomas are the most common indication for hysterectomy. Most tumors are sporadic, but uterine leiomyomas are also a feature of hereditary leiomyomatosis and renal cell cancer syndrome (HLRCC), where biallelic inactivation of *fumarate hydratase* (*FH*) is driving the tumorigenesis.

The overall aim of this thesis project was to take forward the recent finding of *MED12* as a novel driver gene in myomagenesis and to analyze its role in other tumor types. *MED12* exon 2 mutation screening of 1158 samples of various tumor types, both benign and malignant, identified mutations recurrently in uterine leiomyosarcoma (7%) and rarely in colorectal cancer (0.5%). These findings demonstrate that *MED12* exon 2 mutations are not restricted to benign tumors and suggest the possibility that some leiomyosarcomas develop via benign leiomyoma precursors. A sample series representing all the tumor types where *MED12* exon 2 mutations has previously been reported and including only exon 2 mutation-negative tumors were screened for exon 1 mutations. Small in-frame deletions were observed only in conventional uterine leiomyomas with a frequency of 6%, further emphasizing the role of *MED12* in leiomyomagenesis. Gene expression analysis, immunoprecipitations, and kinase activity assays demonstrated that mutations in exon 1 lead to similar global expression profiles and mechanistic effects, including diminished interaction between MED12 and Cyclin C-CDK8/19 and decreased Mediator-associated kinase activity, as previously observed with exon 2 mutations.

Immunohistochemical analysis of the *FH* status and *MED12* exon 1/2 mutation screening confirmed that biallelic *FH* inactivation and *MED12* mutations are mutually

exclusive, both within HLRCC syndrome-associated and sporadic uterine leiomyomas. Based on gene expression profiling, FH-deficient tumors clustered together, whereas HLRCC patients' *MED12* mutation-positive tumors clustered with their sporadic counterparts. These results show that there are at least two distinct molecular mechanisms behind the development of uterine leiomyomas in HLRCC patients: biallelic *FH* inactivation and somatic *MED12* mutations.

A systematic database search utilizing the COSMIC (Catalogue of Somatic Mutations in Cancer) database identified few *MED12* mutations affecting the myoma-linked mutation hotspots also in chronic lymphocytic leukemia (CLL), the most common form of leukemia in adults. Mutation screening of more than 700 CLL samples revealed *MED12* exon 1 and 2 mutations with a frequency of 5% in CLL, making it the first extrauterine cancer where these specific mutations have been observed at a significant frequency. Analysis of molecular and clinical data of 260 patients revealed that positive *MED12* mutation status associated with unmutated status of immunoglobulin heavy chain variable (IGHV) genes and elevated expression of 70 kD zeta-associated protein (ZAP-70), both of which are well-characterized markers of poor prognosis in CLL.

Functional analysis of the first nonsense mutation identified at the amino (N) terminus of MED12 showed that mutant messenger ribonucleic acid (mRNA) escapes nonsense mediated decay (NMD) and produces an N-terminally truncated protein product. Immunofluorescence staining revealed that the mutation prevents the protein's nuclear localization, and, in line with this observation, affinity purification mass spectrometry (AP-MS) analysis showed lost interactions between the mutant protein and other Mediator complex components. A nuclear localization signal (NLS) occurring at the highly conserved N-terminal region of MED12 was predicted and functionally validated. These results suggest an important role for MED12 in normal cell functions and demonstrate that NMD caused by nonsense mutations in early exons can be avoided also in the somatic context.

The results of this study further strengthen the role of *MED12* in pathogenesis of uterine leiomyomas. These findings demonstrate that *MED12* mutations are not restricted to benign hormone-dependent solid tumors, but can be found also in malignant tumors and in hematological diseases. Our results also provide new knowledge about the structure and normal functions of the MED12 protein.

REVIEW OF THE LITERATURE

1. Tumorigenesis

Development and functions of the human body, every organ and tissue, are conducted through the strictly controlled interplay of individual cells in the organism. Cells are regulated by the protein and ribonucleic acid (RNA) products encoded based on the information stored in the genome or by signals coming from their immediate surroundings and environment. Errors in the genetic code, either inherited or somatically acquired mutations, can alter some of the main characteristics of the cell. A subset of these changes can give the cell a growth advantage or ability to evade programmed cell death (apoptosis), leading to clonal expansion. When these kinds of alterations accumulate in a single cell, the formation of a tumor can be initiated (Hanahan and Weinberg, 2000). The vast majority of mutations occur in the non-coding and non-regulatory region of the genome and thus are neither advantageous nor harmful for the cell (Khurana *et al.*, 2016). Some of the errors can be repaired by the cells' own repair mechanisms or cells with detrimental properties can be eliminated. Changes promoting tumorigenesis do, however, occur and accumulate up to a level that only in Finland over 30,000 cancer cases are diagnosed every year (Finnish Cancer Registry, statistics).

The specific capabilities that the neoplastic cells themselves have to reach, or harness the adjacent cells to provide them with, have been postulated for a malignant tumor development. These 'hallmarks of cancer' include sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and ultimately activating invasion and metastasis (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011). In addition, reprogramming of cellular energetics and avoiding immunological destruction are recognized as emerging hallmarks relevant in many, if not all, cancers. Genomic instability and tumor-promoting inflammation are described as characteristics enabling the acquisition of these hallmarks (Hanahan and Weinberg, 2000; Hanahan and Weinberg, 2011). Benign tumors differ from malignant cancer only in their incapability to invade surrounding tissues and metastasize to different sites in the body. Although benign tumors remain within the originating tissue, they can reach a detrimental size or be located in a manner that may cause severe symptoms to the patient (Arafah and Nasrallah, 2001; Stewart, 2015). Generally benign tumors do not dedifferentiate from the originating tissue at similar scale as seen with malignant tumors, and thus they may exert the original function of the tissue, for instance hormone secretion, at abnormal and potentially harmful levels (Arafah and Nasrallah, 2001).

In tumor tissue, cancer cells thrive in association with various different cell types, such as endothelial cells, infiltrating immune cells, and cancer-associated fibroblasts. Non-malignant cells together with transformed tumor cells constitute the tumor

microenvironment that co-evolves and contributes to the progression and growth of the primary tumor and its ability to invade and resist treatment (Hanahan and Coussens, 2012; Junttila and de Sauvage, 2013). Varying composition of the cells, partly due to the environmental differences between distinct regions of the lesion, creates phenotypic and functional intratumoral heterogeneity. These are also affected by genetic heterogeneity, which refers to the different genomic composition of the tumor cells in distinct subclones within the same lesion. Genetic and epigenetic changes as well as heterotypic signaling between cancer cells and the microenvironment occur both in spatial and temporal contexts, leading to intratumoral heterogeneity and tumor evolution. The resulting plasticity allows tumors to adapt to differing environmental pressures, fostering, for instance, tumor growth, dissemination, and therapeutic resistance (Junttila and de Sauvage, 2013; McGranahan and Swanton, 2015). A proposed model of branched tumor evolution, in which subclones are separated by branching and diverge from each other through the acquisition of novel mutations, better explains intratumoral heterogeneity than the conventional linear model with sequential mutation accumulation to a single cell (Swanton, 2012).

2. Genetics of cancer

Mutations in the genome occur spontaneously in every cell cycle with a context-dependent mutation rate that varies throughout the genome. Mutation rate correlates with characteristics such as chromatin organization, gene expression level, and the timing of replication in a way that regions with more heterochromatin-like architecture, low-level expression, and late replication timing have the highest rate (Stamatoyannopoulos *et al.*, 2009; Schuster-Bockler and Lehner, 2012; Lawrence *et al.*, 2013; Watson *et al.*, 2013). Environmental carcinogens, including for example ultraviolet light and radon gas, as well as lifestyle-dependent factors such as smoking and alcohol consumption, cause mutations and increase the risk of the neoplastic transformation of a cell. A mutation providing a selective growth advantage is considered to be a driver mutation, promoting the transformation of the cell into a neoplastic form. Several driver mutations, each adding to the profitable features of the deriving cell clone, are required. Other mutations accumulating in the cell before the neoplastic shift, or along the process, are so-called passenger mutations without any essential contribution to tumorigenesis (Thiagalingam *et al.*, 1996; Greenman *et al.*, 2007; Vogelstein *et al.*, 2013). Based on recent large-scale sequencing studies on different cancers, approximately 200 ‘cancer genes’ with driver mutations have been identified, all involved in central cellular processes: maintenance of genome integrity, regulation of a cell’s fate (differentiation status), and survival (Vogelstein *et al.*, 2013; Vogelstein and Kinzler, 2015). It has been estimated that on average two to eight driver gene mutations are present in a full-fledged malignant tumor, while the vast majority of the somatic variation is composed of passenger mutations (Vogelstein *et al.*, 2013). The acquirement of the needed driver mutations and thus the development from the initiating mutational event to a clinically relevant tumor is a long process that

can last several years or even decades. At the end, the frequency of acquired mutations varies greatly between cancer types, with pediatric cancers and hematological malignancies having the lowest and cancers with strong environmental component, for example lung cancer and melanoma, harboring the highest number of mutations (Lawrence *et al.*, 2013; Vogelstein *et al.*, 2013; Watson *et al.*, 2013). Mutation frequency also varies between patients with the same cancer type, for instance in the case of colorectal cancer where defects in deoxyribonucleic acid (DNA) mismatch repair or replication proofreading lead to a hypermutated state of some tumors (Cancer Genome Atlas Network, 2012; Lawrence *et al.*, 2013; Palles *et al.*, 2013).

2.1 Genomic alterations

A broad spectrum of mutations and aberrations can affect the genes or their regulatory regions relevant in tumor progression. The majority of the alterations observed in protein-coding genes of tumors are single nucleotide variations (SNV). Synonymous mutations do not change the amino acid encoded by the affected codon (nucleotide triplet), whereas missense or nonsense mutations result in a substitution of an amino acid or in a termination codon, respectively. Splice-site mutations affect the processing of the precursor messenger RNA (mRNA) leading to aberrant splicing, for example in the form of exon skipping or intron retention (Jung *et al.*, 2015; Sveen *et al.*, 2016). Synonymous mutations, although harmless in regard to the amino acid sequence, can also cause defects in splicing. These may also affect mRNA stability and translation speed as well as modify microRNA (miRNA) binding sites in coding and untranslated regions (Diederichs *et al.*, 2016). Small scale insertions (ins) and deletions (del), less than 50 base pairs in length, are also frequent in tumors. Changes that result in a varying number of complete codons and thus alter the amino acid sequence are called in-frame mutations. Insertions and deletions which alter the reading frame (frameshift mutations) usually lead to a premature termination codon and nonsense mediated decay (NMD) or a truncated protein product. When occurring in regulatory regions (promoters, enhancers, silencers, insulators), these changes can affect the transcription by altering the transcription factor binding sites or interactions between regulatory elements (Diederichs *et al.*, 2016; Khurana *et al.*, 2016).

Chromosome-level alterations include broader genomic rearrangements such as deletions, amplifications, translocations, insertions, and inversions. These aberrations can affect genes at the exon level or involve one or multiple genes or even larger chromosomal areas. Such rearrangements, as well as chromosomal aneuploidy (changes in the chromosome number), are frequently observed in cancer (Vogelstein and Kinzler, 2004; Negrini *et al.*, 2010; Vogelstein *et al.*, 2013). Chromotripsis is a highly complex manifestation of the chromosomal instability, where one or a few chromosomes are shattered into small fragments and subsequently reassembled randomly (Stephens *et al.*, 2011). This phenomenon is established in tumor development and challenges the conventional conception of the sequential accumulation of genetic alterations (Kloosterman *et al.*, 2014). A high rate of somatic

variation in tumors, generally referred to as genomic instability, can also be caused by microsatellite instability (MSI). Microsatellites are tandemly repeated short nucleotide sequences occurring throughout the genome. In MSI, an impaired DNA mismatch repair system is unable to correct the erroneous number of nucleotide repeats arising during DNA replication (Boland and Goel, 2010). A multitude of mutations can be also induced by defective proof-reading mechanisms of polymerases ε and δ (POLE and POLD1), leading to a hypermutated phenotype as mentioned above (Cancer Genome Atlas Network, 2012; Palles *et al.*, 2013). Excessive accumulation of somatic substitution mutations, mainly cytosine (C) to thymine (T) transitions in the vicinity of chromosomal rearrangements, was observed initially in breast cancer and termed Kataegis (Nik-Zainal *et al.*, 2012). The mutagenic effect of APOBEC DNA cytosine deaminases has been implicated in the phenomenon, although the exact underlying mechanism is yet unclear (Burns *et al.*, 2013).

2.2 Epigenetic and other changes in tumorigenesis

Other factors affecting the functions and regulation of cancer genes, mainly via altered expression, are epigenetic modifications, non-coding RNAs, and infectious agents. Epigenetic changes do not alter the DNA nucleotide sequence but regulate gene expression through mechanisms acting on the genome such as DNA methylation, histone modification, and chromatin remodeling (Choi and Lee, 2013). Epigenetics has a fundamental role in development and it is becoming more evident also in tumorigenesis. Indeed, it has very recently been observed that some pediatric tumors and hematological malignancies may be driven mainly by epigenetic changes (Lee *et al.*, 2012; Feinberg *et al.*, 2016). Genes and regulatory sequences involved in epigenetics can be classified by their function as modulators, modifiers, and mediators (Feinberg *et al.*, 2016). Epigenetic modifiers are responsible for the actual epigenetic modifications, for instance by attaching the methyl groups to cytosine bases, most commonly in the context of CpG dinucleotides, or introducing or removing methyl and acetyl groups involved in histone modification (Choi and Lee, 2013; Feinberg *et al.*, 2016). Mediators are factors that are regulated by these modifiers and whose altered expression promotes tumorigenesis. Upstream of these two groups are the epigenetic modulators that regulate the activity of the epigenetic machinery in a response to stress signals and, in the context of cancer, push the cell towards a neoplastic form (Feinberg *et al.*, 2016). Similarly as with somatic mutations, epigenetic changes can be identified as driver or passenger changes, with the majority being mere passengers (Pon and Marra, 2015). As an example of a driver change affecting an epigenetic mediator would serve promoter hypermethylation and subsequent inactivation of *BRCA1* which has been recurrently observed in sporadic breast and ovarian carcinomas (Cancer Genome Atlas Research Network, 2011; Zhang and Long, 2015).

Non-coding RNAs, particularly miRNAs and long non-coding RNAs (lncRNA), are implicated in tumorigenesis. MiRNAs are short transcripts which are able to bind to mRNAs of protein coding genes and repress their translation. In cancer, miRNA function can be affected through changes in their coding sequence, expression, or in their respective binding sites (Diederichs *et al.*, 2016; Khurana *et al.*, 2016). Over 200 nucleotides long lncRNAs have several roles in cancer biology, for instance through epigenetic regulation, the p53 pathway, and the modulation of miRNA function (Evans et al 2016). Deregulated expression of lncRNAs has been observed in various cancers, but the exact mechanisms and impact in tumorigenesis largely remain to be unraveled (Diederichs *et al.*, 2016; Khurana *et al.*, 2016). A well-known example of the role of infectious agents in cancer development is the human papillomaviruses (HPV) in cervical cancer (Bosch *et al.*, 2002). Expression of the HPV proteins E6 and E7 in epithelial cells leads to degradation of TP53 and the retinoblastoma proteins RB1, RBL1, and RBL2 and subsequently to continuous cell cycle progression and decreased apoptosis. The resulting genomic instability and enhanced cell proliferation drive the malignant transformation, where effective immune evasion is also essential (Crosbie et al 2013). HPV DNA can be detected in virtually all cervical cancers, and the majority are caused by two high-risk HPVs, HPV types 16 and 18 (Bosch *et al.*, 2002).

2.3 Cancer genes

Genes which most commonly drive tumor progression can be classified into oncogenes and tumor suppressor genes by their function and the change resulting from the mutation. Stability genes facilitating tumor development are also described as their own entity (Weinberg, 1994; Vogelstein and Kinzler, 2004; Croce, 2008).

2.3.1 Oncogenes

Proto-oncogenes encode products that induce cell growth and proliferation or inhibit signals leading to cell-cycle arrest or elimination of a cell. In non-neoplastic cells, their expression is tightly controlled to maintain balanced cell growth within the tissue. As a result of an activating 'gain-of-function' mutation, a gene's normal function is accelerated or abnormally activated, or neomorphic activity is acquired, after which the gene is depicted as an oncogene. Typical alterations affecting oncogenes are amino acid substitutions concentrating at specific residues, mutations in the regulatory region, amplifications, and chromosomal rearrangements causing overexpression or producing a novel fusion gene (Weinberg, 1994; Vogelstein and Kinzler, 2004; Croce, 2008). For example, a point mutation affecting one of the critical codons in *RAS* genes (*KRAS*, *HRAS*, and *NRAS* coding for small GTPases) renders them constitutively active oncogenes leading to continuous cell growth (Pylayeva-Gupta *et al.*, 2011). The Philadelphia chromosome is an example of an oncogene activation through formation of a fusion gene. A reciprocal translocation between chromosomes 22 and 9 creates the *BCR-ABL1* fusion gene coding for a constantly active tyrosine kinase.

Rearrangement is observed in certain hematological malignancies, particularly in chronic myeloid leukemia (Advani and Pendergast, 2002). Activating mutations in proto-oncogenes are dominant as a single mutated allele is sufficient to provide the cell with a growth advantage (Weinberg, 1994; Vogelstein and Kinzler, 2004; Croce, 2008).

2.3.2 Tumor suppressor genes

Proteins or RNAs that negatively regulate proliferation or guide the cell to self-destruction ('gatekeepers'), or maintain the cellular microenvironment ('landscapers'), are encoded by tumor suppressor genes (TSG) (Kinzler and Vogelstein, 1997; Kinzler and Vogelstein, 1998). Genetic alterations affecting these genes are usually inactivating 'loss-of-function' mutations that reduce or abolish the function of the gene. These include missense mutations affecting crucial amino acids, nonsense mutations, splice-site mutations, mutations in the regulatory region, subtle intragenic as well as large-scale deletions or insertions, and epigenetic silencing. In order to provide a cell with a selective growth advantage, generally both alleles of the tumor suppressor need to be inactivated (Vogelstein and Kinzler, 2004). This recessive model is characterized by Knudson's two hit hypothesis, and is explicitly relevant in the majority of inherited cancer predisposition syndromes, where the other allele of the TSG is inactivated already in the germ line (Knudson, 1971). In the context of haploinsufficiency, inactivation of only one TSG allele is enough to promote tumorigenesis. In addition, some TSGs display dosage-dependency, where even a modest downregulation of the expression can be sufficient for tumorigenesis in a tissue-specific manner (Alimonti *et al.*, 2010; Berger *et al.*, 2011). For example, only a 20% reduction in *PTEN* expression has been shown to increase tumor susceptibility in mice, most notably the development of mammary tumors in females (Alimonti *et al.*, 2010). A single mutated allele of a tumor suppressor can promote tumorigenesis also via a dominant-negative mode of action. This phenomenon is relevant with tumor suppressors that exert their effect by forming polymeric molecules, for instance with tumor protein p53 (TP53), where the product of the mutated allele is able to sequester also the wild-type product from its normal function (Willis *et al.*, 2004; Payne and Kemp, 2005).

Some genes may function as an oncogene or a tumor suppressor depending on the mutation and the tumor type. An example of this kind of bidirectional function is seen in the case of *Notch (Drosophila) homolog 1 (translocation-associated)* (*NOTCH1*), which, as a result of inactivating mutations, functions as a tumor suppressor in head and neck squamous cell carcinoma (Stransky *et al.*, 2011), and as an activated oncogene in hematological malignancies, such as T-cell acute lymphoblastic leukemia (T-ALL) and chronic lymphocytic leukemia (CLL) (Puente *et al.*, 2011).

2.3.3 Stability genes

Instead of regulating cell proliferation, a class of stability genes, also known as caretaker genes, sustain the genome integrity (Vogelstein and Kinzler, 2004). This class contains genes involved in the repair of nucleotide-level mistakes as well as genes controlling integrity at the chromosomal level, for example the segregation of the chromosomes. Inactivation of stability genes, typically through loss-of-function mutations in both alleles, leads to an increased mutation rate, thus potentiating the cell for tumorigenic changes. Mutations in stability genes are common in inherited tumor susceptibility syndromes. Accordingly, mutations in the mismatch repair genes *MLH1*, *MSH2*, *MSH6*, and *PMS2* underlie hereditary non-polyposis colorectal cancer and endometrial cancer in Lynch syndrome (Sehgal *et al.*, 2014), and mutations in the homologous recombination DNA repair genes *BRCA1* and *BRCA2* predispose to hereditary breast and ovarian cancers (HBOC) (Narod and Foulkes, 2004)

2.4 Inherited susceptibility for cancer

It is estimated that 5 to 10% of all cancer cases arise due to inherited susceptibility for cancer (Nagy *et al.*, 2004). More than a hundred cancer predisposing genes (CPG) conferring high or moderate risk when mutated have been identified by to date (Rahman, 2014). Causative variants in these genes are estimated to account for approximately 3% of cancers. Individuals with germline mutations in CPGs may develop multiple primary tumors usually occurring at an earlier age. The inheritance pattern of cancer predisposition is autosomal dominant for the majority of CPGs, meaning that one affected allele is sufficient to cause the predisposition. Some genes, *BRCA2* and *MLH1*, among others, predispose to pediatric tumors when both alleles harbor germline mutations, whereas monoallelic mutation carriers are prone to adulthood cancers after somatic inactivation of the wild-type allele (Wang *et al.*, 1999; Howlett *et al.*, 2002; Rahman and Scott, 2007). Most of the CPGs are affected by loss-of-function mutations and act as tumor suppressors and specifically as stability genes (Rahman, 2014). Inactivating mutation in the germ line is usually a point mutation or a small-scale insertion or deletion. Inactivation of the wild-type allele in the tumor is most commonly conferred by loss of heterozygosity (LOH), which can be acquired through deletion of the region or copied and replaced from the homologous, mutation containing chromosome (copy-neutral LOH). Different mutations in one gene can confer different cancers. Activating promoter mutations in telomerase encoding *TERT* predisposes carriers to melanoma, whereas exonic mutations cause dyskeratosis congenita syndrome including predisposition for example to head and neck cancer and acute myeloid leukemia (AML) (Nelson and Bertuch, 2012; Horn *et al.*, 2013). In addition, the cancer risk and the clinical phenotype can be modified by other genetic and non-genetic factors (Levy-Lahad and Friedman, 2007; Antoniou *et al.*, 2008; Rahman, 2014; Jori *et al.*, 2015). Approximately half of the CPGs containing high-to-moderate risk variants are recognized to contribute to tumorigenesis also when

harboring only somatic mutations (Rahman, 2014). For instance, somatic alterations affecting the CPGs *TP53* and *RB1* are seen frequently in various cancer types.

In addition to cancers associated with highly penetrant cancer syndromes, 15-20% of all common cancers are considered to be familial. Here, the clustering of certain cancers in a family is observed more often than would be expected by chance (Nagy *et al.*, 2004). Variants conferring low or moderate risk for cancer predisposition, together with shared environmental factors, most likely underlie familial cancer accumulation. Genome-wide association studies including several thousands of cases and controls have identified an increasing number of common polymorphisms as low risk variants in most common cancers (Hosking *et al.*, 2011).

2.5 Genome-wide methods in cancer genetics

Several technical revolutions have been seen in the field of molecular genetics during recent years. In cancer genetics, research has shifted from studying single or a limited number of genes in a given cancer, to describing mutational landscapes and expression profiles of different tumor types, and in clinical context further to deciphering the genomic composition of a certain tumor in a given patient. Microarrays are utilized in various high-throughput analyses to study for instance gene expression, copy number variation, DNA methylation, and histone modifications as well as in genotyping and miRNA profiling. Various applications are based on the detection of labeled target binding to a nucleotide probe attached on a solid surface. Traditional nucleotide sequencing methods have been accompanied by massive parallel sequencing applications, where a large amount of the genome is sequenced simultaneously. In whole exome sequencing (WES) virtually all of the protein-coding genome is captured for sequencing. Sequencing can alternatively be targeted to more restricted scope, covering for instance only a panel of cancer associated genes. In whole genome sequencing (WGS) both the coding (~1.5%) and the non-coding (~98.5%) part of the genome is sequenced. The DNA sequence is fragmented, ligated with synthetic adapters, amplified, and sequenced in parallel. These sensitive, high-throughput next-generation sequencing (NGS) technologies have resulted in an unprecedented ability to study genomic variation in tumors. As methods are evolving and becoming more feasible and inexpensive, they are becoming more common not only in research settings but also in clinical diagnostics. By the year 2015, over 10,000 cancer exomes and more than 2,500 cancer genomes had been sequenced by individual research groups or collaborative sequencing programs such as The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) (Martincorena and Campbell, 2015). These efforts are generating abundant data about somatic changes and mutational processes involved in the tumorigenesis of specific tumor types. Acquired information on cancer-related mutations and rearrangements in individual samples are repositoried in public databases, for instance in COSMIC (Forbes *et al.*, 2008) and the CBioPortal for Cancer Genomics (Gao *et al.*, 2013).

3. Uterine leiomyomas

Uterine leiomyomas, or uterine fibroids, represent one of the most common tumor types in women. The estimates of their overall prevalence vary form 20% to over 70% by the age of 50, and they are clinically relevant in one quarter of women (Cramer and Patel, 1990; Day Baird *et al.*, 2003). Uterine leiomyomas originate from the smooth muscle cells of the myometrium and present as benign tumors with a considerable portion of extracellular matrix (ECM). Lesions can occur in different sections of the uterine wall as submucous (intracavitary), subserous, or intramural leiomyomas (Figure 1), and they may also represent a mixed type in regards to their location (Bajekal and Li, 2000; Stewart *et al.*, 2016). Tumors vary greatly in size (up to 20 cm in diameter), and they can be present as individual or typically as multiple lesions within the uterus (Cramer and Patel, 1990; Day Baird *et al.*, 2003; Heinonen *et al.*, 2014). The spectrum of symptoms caused by these common tumors include for instance abnormal uterine bleeding, discomfort and pain in the pelvic and abdominal areas, and urinary incontinence (Stewart *et al.*, 2016). Uterine leiomyomas may also interfere with the implantation of the embryo and cause complications during pregnancy or labor, especially in the context of submucous leiomyomas (Bajekal and Li, 2000; Pritts *et al.*, 2009). The severity of the symptoms depends on the number, location, and size of the leiomyomas. Because of the notable morbidity to women and the lack of highly specific and long-lasting medical treatments, hysterectomy remains the most common option to deal with the symptoms caused by uterine leiomyomas. Myomectomy or uterine artery embolization are preferred among younger women who wish to maintain their reproductive capability. Medication is applied mainly as a short-term treatment to scale down the symptoms as well as to reduce the tumor size before the surgical operation (Stewart, 2015). Solely direct medical costs caused annually by uterine leiomyomas have been estimated to be $4.1-9.4 billion in the US alone (Cardozo *et al.*, 2012).

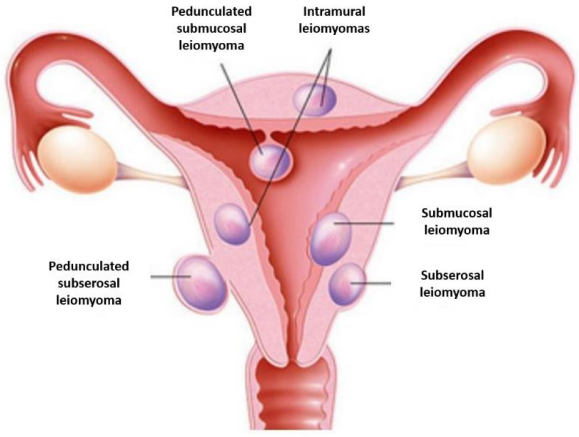

Figure 1. The main types of uterine leiomyomas based on their distinct locations in the uterus. Figure was retrieved August 8th from the NHS Choices web site (http://www.nhs.uk/ Conditions/ Fibroids /Pages /Introduction. aspx) and is reproduced with the permission of the copyright holder.

3.1 Development of uterine leiomyomas

The development and growth of uterine leiomyomas are strongly dependent on the female hormones estrogen and progesterone (recently reviewed in Moravek *et al.*, 2015). In accordance, lesions are most prevalent during the reproductive age and usually shrink along with menopause (Parazzini *et al.*, 1988; Cramer and Patel, 1990; Flake *et al.*, 2003). Nulliparity and early onset of regular menstrual cycles are risk factors for uterine leiomyomas, while multiparity and later menarche have been associated with reduced risk (Parazzini *et al.*, 1988; Parazzini *et al.*, 1996; Marshall *et al.*, 1998; Sato *et al.*, 2000; Faerstein *et al.*, 2001; Terry *et al.*, 2010). Also, a high body mass index and hypertension have been suggested as risk factors for developing leiomyomas (Summers *et al.*, 1971; Ross *et al.*, 1986; Shikora *et al.*, 1991; Faerstein *et al.*, 2001; Wise *et al.*, 2005). Inherited genetic factors contribute to the risk as well, since women whose first-degree relatives are diagnosed with uterine leiomyomas have elevated risk of developing leiomyomas (Vikhlyaeva *et al.*, 1995; Sato *et al.*, 2002). In addition, ethnicity strongly influences the probability of leiomyomas. Among African-American women, uterine leiomyomas are more common (cumulative incidence by age 50 >80% compared to ~70% in Caucasian women), develop earlier, and furthermore appear to be more symptomatic than in Caucasian women (Kjerulff *et al.*, 1996; Faerstein *et al.*, 2001; Day Baird *et al.*, 2003).

Traditionally, uterine leiomyomas have been considered as monoclonal tumors originating from one single cell (Linder and Gartler, 1965; Townsend *et al.*, 1970; Zhang *et al.*, 2006). In some cases, multiple tumors within a single uterus have been observed with shared clonal origin, demonstrating a possibility of leiomyoma dissemination (Nilbert *et al.*, 1990; Canevari *et al.*, 2005; Mehine *et al.*, 2013; Mehine *et al.*, 2015). The exact origins and mechanisms of initiation are still unresolved issues in the development of leiomyomas. Recent studies identifying stem cell-resembling cell populations in the context of both the myometrium and leiomyoma have postulated a hypothesis whereby each lesion originates from a single transformed myometrial stem cell and possibly gains further growth advantage through additional genetic changes occurring in the developing tumor (Ono *et al.*, 2007; Ono *et al.*, 2014a).

In addition to conventional uterine leiomyomas, which account for ~90% of all tumors, several histopathological subtypes with distinct morphological features or unusual growth patterns are recognized (Oliva *et al.*, 2014). Some variants, for instance mitotically active, with bizarre nuclei (formerly called atypical), cellular and highly cellular uterine leiomyomas present histopathologically different characteristics that resemble malignancy, yet as a whole, these tumors are considered benign (Oliva *et al.*, 2014). True, and moreover highly malignant, counterpart of uterine leiomyoma is leiomyosarcoma of the myometrium. These tumors usually display morphological features that are seen in distinct uterine leiomyoma subtypes (increased mitotic activity, nuclear atypia, and high cellularity) as well as tumor cell necrosis (Oliva *et al.*, 2014). Tumor type is relatively rare, with an annual incidence

of less than 0.4 cases/100,000 women, but unfortunately it is usually diagnosed in pathological examination only after the appearance of surgery-requiring severe symptoms (Toro *et al.*, 2006; Koivisto-Korander *et al.*, 2012).

3.2 Chromosomal alterations in uterine leiomyomas

Nearly half of uterine leiomyomas present different chromosomal aberrations (Nibert and Heim, 1990; Vanni *et al.*, 1991; Rein *et al.*, 1991; Sandberg, 2005a). Recurrent non-random rearrangements affect for example the *high-mobility group AT-hook 1* and *2* genes at 6q21 (*HMGA1*) and at 12q15 (*HMGA2*), accounting for ~5% and ~20% of cytogenetically aberrant lesions, respectively (Meloni *et al.*, 1992; Van de Ven, 1998; Ligon and Morton, 2000). The *HMGA2* locus-containing region at 12q13-15 is recurrently altered also in other benign tumors of mesenchymal origin, such as lipomas, endometrial polyps, and breast fibroadenomas (Vanni *et al.*, 1993; Schoenmakers *et al.*, 1994; Schoenmakers *et al.*, 1995; Dal Cin *et al.*, 1995; Staats *et al.*, 1996; Sandberg, 2005a). In uterine leiomyomas, the most commonly observed rearrangement is the reciprocal translocation between chromosomes 12 and 14, t(12;14)(q14-q15;q23-q24), where *RAD51 homolog B* (*RAD51B*) at 14q23-24 constitutes the translocation partner for *HMGA2* (Vanni *et al.*, 1991; Rein *et al.*, 1991; Ingraham *et al.*, 1999; Schoenmakers *et al.*, 1999; Quade *et al.*, 2003). This, as well as other rearrangements affecting the *HMGA2* locus or the region upstream to it, usually lead to overexpression of the gene (Gattas *et al.*, 1999; Tallini *et al.*, 2000; Quade *et al.*, 2003). *HMGA2* is also overexpressed in uterine leiomyomas without aberrations affecting 12q14-15, possibly through altered functions of the *HMGA2*-repressive miRNA let-7 (Mayr *et al.*, 2007; Peng *et al.*, 2008; Klemke *et al.*, 2009). Overexpression of HMGA proteins is associated with a highly malignant phenotype in several malignant tumor types (Pallante *et al.*, 2015). HMGA proteins modify the chromatin conformation and thus function as indirect regulators of transcription (Fusco and Fedele, 2007), whereas RAD51B is a DNA repair protein involved in homologous recombination of double-strand DNA breaks (Albala *et al.*, 1997; Suwaki *et al.*, 2011).

Other recurrently observed cytogenetic aberrations in uterine leiomyomas are the trisomy of chromosome 12 and rearrangements affecting the long arm (q) of chromosome 7 (Boghosian *et al.*, 1988; Vanni *et al.*, 1989; Vanni *et al.*, 1992; Sandberg, 2005a). The minimal region commonly deleted at 7q has been mapped to 7q22, wherein *cutlike homeobox 1* (*CUX1*) has been identified as the most likely target gene affected by various rearrangements (Zeng *et al.*, 1997; Schoenmakers *et al.*, 2013; Mehine *et al.*, 2013). The CUX1 transcription factor regulates cell proliferation and invasion, among other cellular processes (Hulea and Nepveu, 2012), and in addition to its putative role in uterine leiomyomas it has been implicated as a haploinsufficient tumor suppressor in myeloid cancers (McNerney *et al.*, 2013; Wong *et al.*, 2014). In a small number of uterine leiomyomas alterations affecting collagen component-coding genes *collagen type IV alpha 5* and *alpha 6* (*COL4A5* and

COL4A6) on the X chromosome have been identified (Mehine *et al.*, 2013). Germline alterations in these genes lead to a rare condition known as Alport syndrome and diffuse leiomyomatosis, predisposing patients to leiomyomas of various organs (Garcia-Torres *et al.*, 2000; Thielen *et al.*, 2003). Some of the above-mentioned chromosomal rearrangements, as well as less frequent alterations in uterine leiomyomas affecting chromosomes 1, 3, 10, 13, and X, may co-occur with the more prevalent cytogenetic changes (Sandberg, 2005a; Mehine *et al.*, 2013). Alterations occurring in addition to the primary tumorigenic events are considered to be secondary changes which are not required but are advantageous for the tumor.

3.3 Hereditary leiomyomatosis and renal cell cancer syndrome

Biallelic inactivation of *fumarate hydratase* (*FH*) is additional, although very rare, event underlying the development of uterine leiomyomas. This phenomenon usually occurs in the context of a tumor predisposition syndrome, hereditary leiomyomatosis and renal cell cancer (HLRCC, Mendelian Inheritance in Man [MIM] database 150800). HLRCC is an autosomal dominant syndrome where heterozygous germline mutations in the *FH* gene predispose mutation carriers to cutaneous and uterine leiomyomas and in some families also to aggressive renal cell cancer (Kiuru *et al.*, 2001; Launonen *et al.*, 2001; Alam *et al.*, 2001; Tomlinson *et al.*, 2002). *FH* locates to 1q43 and encodes the fumarate hydratase enzyme. In its homotetramer form, it functions as a catalyst in tricarboxylic acid cycle (TCAC) hydrating fumarate to malate. Germline mutations are mainly missense and nonsense point mutations occurring throughout the gene. Frameshift mutations, splice-site mutations, and deletions removing the entire gene or part of it have also been observed (Alam *et al.*, 2005b; Wei *et al.*, 2006; Bayley *et al.*, 2008; Ahvenainen *et al.*, 2008; Gardie *et al.*, 2011; Smit *et al.*, 2011). *FH* is a tumor suppressor, and, accordingly, biallelic inactivation of the gene is observed in syndrome-associated tumors. The wild-type allele is most commonly lost through LOH (Kiuru *et al.*, 2001; Launonen *et al.*, 2001; Tomlinson *et al.*, 2002; Alam *et al.*, 2003). Somatic point mutations are also observed as a second hit, and occasionally both alleles are inactivated through somatic events (Kiuru *et al.*, 2002; Lehtonen *et al.*, 2004; Harrison *et al.*, 2015). Homozygous or compound heterozygous germline mutations in *FH* cause fumarase deficiency syndrome (FHD, MIM 606812) with severe neurological impairment that usually leads to death in early childhood (Zinn *et al.*, 1986; Bourgeron *et al.*, 1994).

Cutaneous leiomyomas (piloleiomyomas) are the most prevalent feature of HLRCC syndrome. Uterine leiomyomas are present in the great majority of women with HLRCC (Toro *et al.*, 2003; Alam *et al.*, 2005a; Wei *et al.*, 2006; Stewart *et al.*, 2008). Syndrome-associated FH-deficient uterine leiomyomas are characterized by particular clinical and morphological features compared to sporadic leiomyomas. Tumors are usually more numerous and they are diagnosed earlier in HLRCC patients. In general, they are more symptomatic than their sporadic counterparts and thus require surgical treatment more often and at a younger age (Toro *et al.*, 2003; Alam *et al.*, 2005a; Wei

et al., 2006; Stewart *et al.*, 2008; Lehtonen, 2011). HLRCC-associated uterine leiomyomas represent characteristics of histopathological subtypes, typically those of leiomyomas with bizarre nuclei or cellular leiomyomas, whereas atypical mitoses and tumor cell necrosis, defining malignancy, are not observed. In addition, cell density and organization of the ECM is observed to be higher in these tumors. Similar nuclear features as described in syndrome-associated renal cell cancers are also present in uterine leiomyoma cells, including prominent eosinophilic nucleoli and clear haloes surrounding them, and an ovoid shape of the nuclei (Launonen *et al.*, 2001; Sanz-Ortega *et al.*, 2013; Joseph *et al.*, 2015). Characteristics vary, however, to the extent that identification of syndrome-associated leiomyomas is not feasible solely based on morphological criteria (Alsolami *et al.*, 2014). Renal cell cancer is observed only in a subset of HLRCC families, and more frequent occurrence has been noted in families from Finland and the US. HLRCC-associated renal cell cancers develop at a younger age, are highly aggressive, and represent usually type 2 papillary histology (Launonen *et al.*, 2001; Tomlinson *et al.*, 2002; Toro *et al.*, 2003; Grubb *et al.*, 2007; Vahteristo *et al.*, 2010).

The mechanisms underlying the pathogenesis of FH deficiency are not yet fully understood. One possible explanation may involve the stabilization of hypoxia-inducible factor 1 alpha (HIF1A) by the accumulation of fumarate and subsequent activation of growth promoting hypoxia pathways (Isaacs *et al.*, 2005; Pollard *et al.*, 2005; Pollard *et al.*, 2007; Koivunen *et al.*, 2007). Despite consistent stabilization of HIF1A in FH-deficient tissues, it has been observed that in the Fh1-deficient mice, formation of renal cysts is not dependent on Hif activities (Adam *et al.*, 2011). An alternative tumorigenic route has been suggested, in which the accumulation of fumarate leads to succination of critical cysteine residues in KEAP1, stabilization of NRF2, and subsequent activation of the antioxidant response pathway (Adam *et al.*, 2011; Ooi *et al.*, 2011; Taguchi *et al.*, 2011). Succination, a stable modification of cysteine residues to S-(2-succinyl)-cysteine (2SC), is a direct and specific consequence of elevated fumarate levels in the cell and can be robustly detected with an antibody raised against it (Nagai *et al.*, 2007; Frizzell *et al.*, 2011; Bardella *et al.*, 2011).

3.4 *MED12* mutations in uterine leiomyomas

At the beginning of the next-generation sequencing era, WES was applied to 18 uterine leiomyomas and corresponding normal myometrium tissue samples from 17 Finnish patients. Analysis identified ten tumors with a somatic mutation in the *mediator complex subunit 12* (*MED12*) gene (Mäkinen *et al.*, 2011b). A validation set of 207 uterine leiomyomas from 63 individual patients showed similar occurrence of somatic *MED12* mutations, yielding a mutation frequency as high as 71% (159/225). All mutations were located at a very specific region in the gene comprising the end of the first intron and the beginning of the second exon. Nearly 70% of observed mutations were single nucleotide substitutions affecting codon 44, coding for a glycine amino

acid (G). Most of these mutations resulted in a replacement of this glycine residue with an aspartic acid residue (D), while all other possible amino acid substitutions were observed as well. The rest of the mutations were other missense mutations or small frame-retaining insertions and deletions. Codons 36 (leucine, L) and 43 (glutamine, Q) as well as an intronic T nucleotide eight base pairs before the intron-exon boundary formed additional mutation hotspots, albeit with substantially lower mutation frequencies. Mutations in any other part of the gene were not observed by WES nor by direct Sanger sequencing of 20 *MED12* mutation-negative and 10 *MED12* mutation-positive samples. All observed mutations were heterozygous and complementary deoxyribonucleic acid (cDNA) sequencing of 16 *MED12* mutation-harboring tumors showed that the mutant allele of the X-chromosomal *MED12* was predominantly expressed in all the tumors. In addition, cDNA sequencing verified the splice site effect of the intronic variant at nucleotide position c.100-8 as the six last intronic nucleotides were added to the transcript (Mäkinen *et al.*, 2011b).

Frequent *MED12* mutations in uterine leiomyomas were studied also in a population with different ethnic backgrounds. Sanger sequencing-based analysis included 28 uterine leiomyomas from 18 South African women: six black African women and 12 women with mixed ancestry (colored) (Mäkinen *et al.*, 2011a). Similar kinds of mutations affecting the same region of *MED12* were observed in 50% of the tumors. The total mutation frequency was lower than that observed in uterine leiomyomas derived from Caucasian patients, but the difference leveled out when comparing tumors with a diameter ≥5.5 cm, which are predominant in women with African ancestry. Instead, a statistical difference in the mutation frequencies was observed when only smaller tumors (<5.5 cm) were analyzed (Mäkinen *et al.*, 2011a). These compelling findings were soon validated also by the WES analysis of uterine leiomyomas from North American women. In this study, *MED12* mutations were observed with a frequency of 68%. This study comprised both black and white American women, and here the *MED12* mutation frequency was slightly higher among uterine leiomyomas from black American women, however not reaching statistical significance (McGuire *et al.*, 2012). Again, all the observed mutations were missense mutations affecting the mutational hotspots identified in the Finnish study, or alternatively insertion/deletion mutations that did not change the reading frame.

4. MED12

MED12 is a large gene consisting of 45 coding exons spanning a genomic area of 24 kb. It is located on the long arm of the X chromosome, at Xq13 (X: 71,118,556 - 71,142,454) (Philibert *et al.*, 1998; Philibert *et al.*, 1999; Ensembl database: Homo sapiens, genome assembly GRCh38.p7, MED12). As an X-chromosomal gene, *MED12* is subjected to lyonization, a random inactivation of one of the two X-chromosomes in females. This process balances the expression of X-linked genes between males and females, and it is irreversible in the particular cell and its descendants (Lyon, 1961; Gartler and Goldman, 2001). As are most of the Mediator

subunits, *MED12* is conserved among eukaryotes, and in humans it is more conserved than average genes (Borggrefe *et al.*, 2002; Kitano *et al.*, 2003; Bourbon *et al.*, 2004; Bourbon, 2008). The gene encodes a transcript of 6,795 base pairs, which is translated into the MED12 protein consisting of 2177 amino acids, with a molecular mass of 243,081 Da (UniProt database: Q93074, MED12_Human; UniProt Consortium, 2015). The amino acid sequence of MED12 is largely unique, and it is traditionally divided into four domains characterized by an enrichment of certain amino acids: a leucine-rich L domain (aa 1-500), a leucine-serine-rich LS domain (aa 501-1650), a proline-glutamine-leucine-rich PQL domain (aa 1651-2086), and a glutamine-rich opposite paired (OPA) domain (aa 2087-2177) (Zhou *et al.*, 2002) (Figure 2A). The Med12-PQL domain is listed in the protein families (Pfam) database and is described as Eukaryotic Mediator 12 catenin binding site at amino acids 1819-2022 in MED12 (Pfam database, MED12_Human; Finn *et al.*, 2016). In the database, the Med12 and Med12-LCEWAV domains are also characterized in the human MED12 protein, at amino acids 104-161 and 286-757, respectively (Figure 2). Any specific functions of these domains have not been described, but they contain sequence motifs that are conserved in eukaryotes. MED12 is expressed at comparatively similar levels throughout a range of human tissues and in all developmental stages (Philibert *et al.*, 1999; Ito *et al.*, 1999).

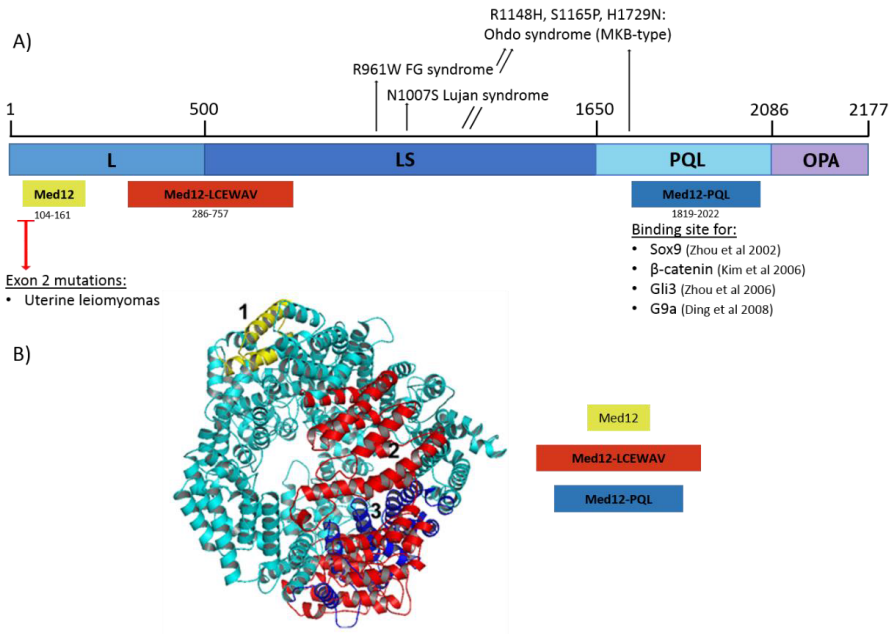

Figure 2. Schematic representation of the human MED12 protein. A) The amino acid sequence is divided into four domains according to their prevailing amino acid content (L=leucine-rich; LS=leucine-serine-rich; PQL=proline-glutamine-leucine-rich; OPA=opposite paired domain). Domains listed in the Pfam database for MED12 are marked under the amino acid domain alignment. Binding sites for various interactors are recognized in the PQL domain. The first germline mutations identified in *MED12*-related XLID syndromes (FG, Lujan-Fryns, OSMKB) are marked above the alignment. Frequent somatic exon 2 mutations observed in uterine leiomyomas are indicated with a red arrow. **B)** The predicted 3D structure of MED12. Conserved domains identified in the protein are indicated as yellow (Med12), red (Med12-LCEWAV), and blue (Med12-PQL) ribbons. The 3D-structure is from the article of Banaganapalli *et al*., 2016, and is reproduced with the permission of the copyright holder

4.1 MED12 as a subunit of the Mediator complex

MED12 is a subunit of the Mediator, a large multiprotein complex exceeding a molecular weight of 1.8 MDa when containing all associated modules (Taatjes 2010). Mediator is a central regulator of RNA polymerase II (Pol II)-dependent transcription by forming a physical connection between regulatory elements and the transcription machinery. In addition to transmitting regulatory signals from DNA-bound transcription factors to Pol II, Mediator is involved in various steps along the transcription process, such as the organization of chromatin structure and in transcription initiation and elongation (Malik and Roeder, 2010; Allen and Taatjes, 2015). The complex was initially discovered in the yeast *Saccharomyces cerevisiae* and has since been observed to exist with a highly conserved composition throughout eukaryotes (Myers and Kornberg, 2000; Bourbon, 2008; Tsai *et al.*, 2013; Tsai *et al.*, 2014; Allen and Taatjes, 2015). In higher eukaryotes, Mediator is composed of 30 subunits, which are divided into four distinct modules according to their location and interactions within the complex. The core of the complex is formed by the 'head', 'middle', and 'tail' modules and constitute the active holoenzyme by binding the Pol II

enzyme (Asturias *et al.*, 1999; Näär *et al.*, 2002; Tsai *et al.*, 2014). The fourth module, associating reversibly with the Mediator core, is the 'CDK8 kinase module' comprising Mediator subunit MED12, Mediator complex subunit 13 (MED13), Cyclin C (CCNC), and Cyclin-dependent kinase 8 or 19 (CDK8/19) (Borggrefe *et al.*, 2002; Bourbon, 2008) (Figure 3).

The CDK8 kinase activity of the module has been shown to disrupt the formation of the transcription preinitiation complex by phosphorylating the Pol II carboxy-terminal domain (CTD) as well as to hinder the subsequent transcript initiation by phosphorylating the general transcription initiation factor IIH (Hengartner *et al.*, 1998; Akoulitchev *et al.*, 2000). In addition, when interacting with the core Mediator, the kinase module precludes Pol II binding, and thus, mostly suppressive regulation of transcription has been associated with it (Elmlund *et al.*, 2006; Knuesel *et al.*, 2009a). More recently, the kinase module has been implicated as an activator of transcription in several instances, such as in the p53 network (Donner *et al.*, 2007), in the HIFA mediated hypoxia response (Galbraith *et al.*, 2013), and in wingless-related integration site (Wnt)/*β*-catenin signaling (Kim *et al.*, 2006; Galbraith *et al.*, 2010; Clark *et al.*, 2015). The detailed mechanisms of kinase module-mediated transcription activation are not yet fully understood, but they might involve precise scheduling of Pol II/Mediator interactions and kinase module recruitment or alternatively rely on structural changes in the Mediator itself (Galbraith *et al.*, 2013; Tsai *et al.*, 2014; Wang *et al.*, 2014b). Electron microscopic imaging and immunochemical analyses showed MED13 to be the main link between the kinase module and the Mediator core, especially the middle module of the Mediator. In the kinase module, MED12 is in the center of the module, connecting the Cyclin C/CDK8 pair to MED13 (Tsai *et al.*, 2013). MED12 has been shown to be required for the kinase activity of CDK8 in higher eukaryotes. While the majority of the kinase modules function in association with the Mediator, a proportion exist as discrete modules, retaining their kinase activity (Knuesel *et al.*, 2009b).

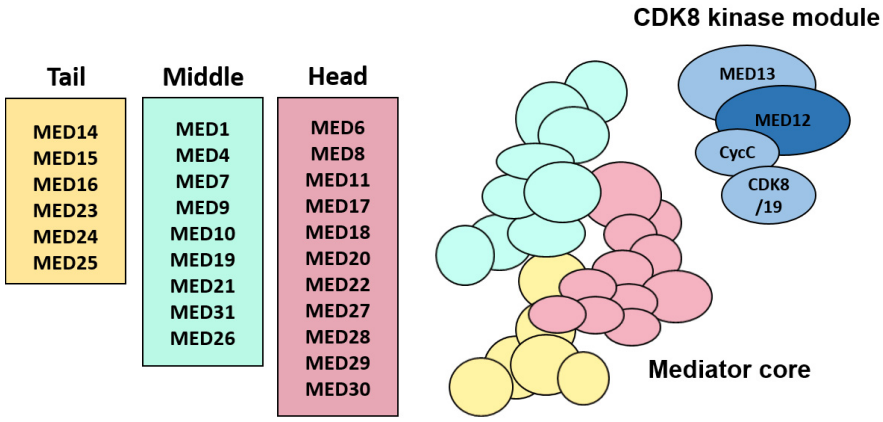

Figure 3. Schematic representation of the Mediator core and CDK8 kinase module. The human Mediator complex comprises 30 subunits. MED12 is a subunit of the CDK8 kinase module, which it forms together with MED13, Cyclin C, and CDK8/19. Division of the Mediator core subunits into head, middle, and tail modules is adapted from the study of Tsai *et al*., 2014.

Paralogs for the kinase module subunits CDK8, MED12, and MED13 have been identified in vertebrates (Sato *et al.*, 2004; Bourbon, 2008; Tsutsui *et al.*, 2008; Clark *et al.*, 2015). MED12 paralog MED12L on chromosome 3 was initially identified as NOPAR (no OPA repeat), with 61% sequence homology but without the OPA domain found in original MED12 (Joensuu *et al.*, 2001). The PQL and OPA domains were recently identified also in the paralog, thereafter named MED12L (Vogl *et al.*, 2013). The same conserved domains (Med12, Med12-LCEWAV, and Med12-PQL) are listed in the Pfam database for both MED12 and MED12L. Of the kinase module subunits, CDK8 paralog CDK19 most closely resembles its counterpart. Their amino acid similarity exceeds 90% at the kinase domain and shows high conservation among vertebrates (Manning *et al.*, 2002; Tsutsui *et al.*, 2008). Overlapping expression-regulating functions have been described to these proteins (e.g. Pol II CTD phosphorylation), while unique targets in context-dependent settings also exist (Tsutsui *et al.*, 2011). Moreover, opposing functions for these subunits have been observed in viral transcription activator VP16-dependent transcription (Tsutsui *et al.*, 2008).

4.2 MED12 in signaling pathways

Only a few functional domains have been described in MED12 (Figure 2). The protein region containing amino acids 1651-2086, covering the Med12-PQL domain, have been identified as a direct binding target for the *β*-catenin transactivation domain (Kim *et al.*, 2006). Through this interaction, MED12 and the Mediator complex are recruited to activate target gene transcription in canonical Wnt/*β*-catenin signaling. The relevance of Med12 to Wnt/*β*-catenin signaling as well as Wnt/planar cell polarity (PCP) signaling was demonstrated when Med12-null and hypomorphic mouse

embryos showed severe defects in the developmental events regulated by these signaling pathways (Rocha *et al.*, 2010).

In addition to the well-established interaction with *β*-catenin, MED12 has been implicated in a few other signaling pathways, with an emphasis on developmental signaling (Yin and Wang, 2014; Clark *et al.*, 2015). The carboxy (C) -terminal area of MED12 containing the PQL domain has been identified as a binding region also for the Sox9 and Sox10 transcription factors, thus linking MED12 for example to the developmental regulation of the nervous system and chondrocyte differentiation in which designated Sox family members are involved (Zhou *et al.*, 2002; Rau *et al.*, 2006; Wang *et al.*, 2006; Vogl *et al.*, 2013). Other members of Sox family, Sox32 and Sox4b, have also been implicated to be regulated by Med12 in the development of the endoderm in Zebrafish (Shin *et al.*, 2008). Furthermore, the MED12 PQL domain has been shown to physically interact with the Sonic hedgehog (Shh) signaling effector glioma-associated oncogene family zinc finger 3 (Gli3), more specifically with its MED12/Mediator binding transactivation domain (Zhou *et al.*, 2006). Interaction between MED12 and activated Gli3 has been demonstrated to lead to Mediator-dependent suppression of Shh target gene transcription (Zhou *et al.*, 2006; Zhou *et al.*, 2012).

MED12 has also been shown to exert repressive actions on transcription through epigenetic regulation. EHMT2 methyltransferase, also known as G9a, binds directly to the MED12 PQL domain and is connected through the Mediator core (more accurately through MED19 and MED26) to the RE1 silencing transcription factor (REST). This complex leads to transcriptional suppression of specific neuronal genes in non-neuronal cells via H3K9 histone di-methylation (Ding *et al.*, 2008; Ding *et al.*, 2009).

4.3 *MED12* mutations in the germ line

The gene coding for MED12 was initially described as a human OPA-containing gene (HOPA) because of the rare polymorphism in its OPA domain. This variant is an insertion of 12 base pairs in exon 43, leading to the addition of an extra QQHQ amino acid stretch to the glutamine-rich sequence of the region. Originally, the insertion was reported to be associated with an X-linked mental retardation and hypothyroidism syndrome (Philibert *et al.*, 1998; Philibert *et al.*, 1999), but instead an association with increased risk for schizophrenia was subsequently demonstrated (Philibert, 2006; Philibert *et al.*, 2007).

Germline mutations in *MED12* have been connected to different X-linked intellectual disability (XLID) syndromes (Figure 2A). Missense mutations leading to substitutions R961W or G958E in the LS domain of MED12 underlie Opitz-Kaveggia syndrome, also named FG syndrome-1 (MIM 305450) (Opitz and Kaveggia, 1974; Risheg *et al.*, 2007; Rump *et al.*, 2011). A nearby missense mutation affecting the same domain and

leading to an N1007S substitution has been identified in Lujan-Fryns syndrome (MIM 309520) (Lujan *et al.*, 1984; Fryns and Buttiens, 1987; Schwartz *et al.*, 2007). These syndromes present a broad variety of clinical manifestations caused by developmental defects, with several overlapping symptoms, such as intellectual disability, malformation of the corpus callosum, macrocephaly, and a tall forehead (Graham and Schwartz, 2013). *MED12* missense mutations causing both Opitz-Kaveggia and Lujan-Fryns syndromes have been shown to disrupt its function in REST-mediated gene silencing as well as in Gli3-dependent Shh signaling (Ding *et al.*, 2008; Zhou *et al.*, 2012).

At least three different *MED12* missense mutations causing amino acid changes at the C-terminal end of MED12 (R1148H, S1165P, and H1729N) have been linked to Maat–Kievit–Brunner type Ohdo syndrome (OSMKB; MIM 300895). This syndrome is an X-linked subtype of blepharophimosis-intellectual disability syndromes with characteristic facial features as well as intellectual and developmental delays (Maat-Kievit *et al.*, 1993; Vulto-van Silfhout *et al.*, 2013). Additional *MED12* missense mutations, and recently also one frameshift mutation, have been linked to X-linked intellectual deficiency with milder syndromic features or with a different composition of manifestations than in previously described syndromes. In some families, intellectual impairment has also been observed in heterozygous females (Lesca *et al.*, 2013; Bouazzi *et al.*, 2015; Langley *et al.*, 2015; Prontera *et al.*, 2016). Phenotypic variability and differing severity of the features between patients, even when linked to same *MED12* mutation, implicate a spectrum of MED12-related disorders rather than separate syndromes (Isidor *et al.*, 2014; Langley *et al.*, 2015).

4.4 Somatic *MED12* mutations in other tumor types?

Uterine leiomyomas are the first tumors where recurrent somatic *MED12* mutations have been observed. The high mutation frequency and specific location of non-truncating mutations further confirmed their essential role in the tumorigenesis of uterine leiomyomas. The smooth muscle cells of the myometrium that serve as progenitors for uterine leiomyomas originate from the mesenchymal tissue. The mesenchyme is formed mainly by the mesodermal germ layer and gives rise to the tissues of the circulatory and lymphatic systems as well as smooth muscles, bones, and cartilage. Another distinctive feature of uterine leiomyomas is their dependence on the steroid hormones estrogen and progesterone. A large proportion of other female tumors, especially those of the breasts, ovaries, and endometrium, are initially dependent on steroid hormones, although later in the development they might gain hormone-independency. Other tumor types deriving from tissues with mesenchymal origin, such as bone and cartilage sarcomas, leiomyosarcomas, leiomyomas of other organs, endometrial polyps, and lipomas, as well as estrogen and progesterone-dependent tumors, might share similar mechanisms in tumorigenesis and harbor *MED12* mutations.

Preliminary gene expression and pathway enrichment analyses performed on *MED12* mutation-positive uterine leiomyomas suggested an involvement of aberrant Wnt signaling in leiomyomagenesis (Mäkinen *et al.*, 2011b). Furthermore, as mentioned above, MED12 has been shown to bind directly with β-catenin and regulate the canonical Wnt/β-catenin signaling pathway (Kim *et al.*, 2006). Dysregulation of Wnt signaling is one of the central mechanisms involved in human cancers. For example, the loss of APC function, both in the context of familial adenomatous polyposis and sporadic colorectal cancers, leads to the accumulation of β-catenin and activation of the Wnt signaling pathway (Polakis, 2012; Novellasdemunt *et al.*, 2015). A role for MED12 has also been suggested in hematopoiesis as it has been observed to regulate the differentiation of certain myeloid lineage cells in *Drosophila* and zebrafish (Gobert *et al.*, 2010; Keightley *et al.*, 2011). In addition, retroviral insertions affecting the *Med12* locus have been identified in mice with leukemia (Dave *et al.*, 2009). After the initial identification of somatic *MED12* mutations, relevant expression data stored in public repositories was analyzed utilizing the Genesapiens database (now Medisapiens). The database contained gene expression data from 175 human tissues (43 normal tissue types, 68 cancer types, and 64 other diseases) collected from nearly 10,000 expression array experiments (Kilpinen *et al.*, 2008). Analysis revealed *MED12* expression to be clearly upregulated in the leukemia subtypes AML and ALL in comparison with other tumor types. Based on these observations, tumors with aberrant Wnt signaling as well as hematological diseases compose yet additional groups of neoplasms where *MED12* mutations might play a role.

AIMS OF THE STUDY

The overall aim of this thesis project was to take forward the recent finding of *MED12* as a novel driver gene in myomagenesis and to analyze its role in other tumor types. The aim was also to investigate the functional impacts of the observed mutations and to shed light on the normal functions of MED12.

The specific aims were as follows:

1. To study the role of *MED12* exon 2 mutations in various tumor types

2. To investigate if mutations also exist in exon 1 of *MED12* and if they share the same functional effect as exon 2 mutations

3. To study the role of *MED12* in the tumorigenesis of HLRCC patients' uterine leiomyomas and to analyze if *MED12* mutations and biallelic *FH* inactivation are mutually exclusive

4. To analyze the mechanistic effects of the unusual *MED12* exon 1 nonsense mutation observed in a patient with T-cell acute lymphoblastic leukemia

MATERIALS AND METHODS

1. Study subjects and samples

1.1 Tumor samples (I-V)

In total, 2259 tumor samples, both benign and malignant, were included in the mutation screenings performed in studies I-IV. The sample series consisted of mesenchymal tumors, estrogen-progesterone dependent tumors, hematological malignancies, and colorectal cancer samples (Table 1).

In Study I, 1158 tumor samples were screened for *MED12* exon 2 mutations. The sample series included uterine leiomyosarcomas and other sarcomas, gastrointestinal stromal tumors, extrauterine leiomyomas, endometrial polyps, lipomas, ovarian cancers, breast cancers, acute myeloid leukemias, acute lymphoblastic leukemias, myeloproliferative neoplasms, and colorectal cancers. The series was collected from several university and central hospitals in Finland, Denmark, and the United States. The samples were obtained either as preserved tumor samples (fresh frozen, FF, or formalin-fixed paraffin-embedded, FFPE) or as ready-extracted DNA. Samples of AML and myeloproliferative neoplasms from Memorial Sloan-Kettering Cancer Center, NY, USA, were extracted and analyzed at the institution's facilities.

In Study II, 611 samples previously identified as *MED12* exon 2 mutation-negative, were screened for mutations in exon 1 of the gene. The sample series consisted of uterine leiomyomas, extrauterine leiomyomas, endometrial polyps, uterine leiomyosarcomas, other sarcomas, and colorectal cancer samples, representing all the tumor types where recurrent *MED12* exon 2 mutations had been observed until then.

Altogether, 188 uterine leiomyoma samples were screened for *MED12* exon 1/2 mutations and their *FH* status was analyzed in Study III. Of these, 122 were from 27 Finnish HLRCC patients representing 11 families (89 FFPE and 33 FF samples) and 66 were sporadic uterine leiomyoma samples (FFPE) from as many individuals. HLRCC families included in the study are listed in Table 2.

Three independent sample series comprising 746 CLL samples were collected for *MED12* exon 1 and 2 mutation screening in Study IV. Two of the series were collected from the United States, from the Tissue Core Biorepository of CLL Research Consortium (CRC; 278 samples) and The Ohio State University's Human Genetics Sample Bank (292 samples). US samples were extracted and sequenced at the respective institutions. The third sample series was collected from the Helsinki University Central Hospital clinical sample collection (176 samples). More detailed

information about the series and the collection procedures can be found in the Supplementary Data of the original publication IV.

In Study V, one T-ALL patient sample harboring a *MED12* nonsense mutation (c.97G>T, p.E33X) (Kontro *et al.*, 2014) served as the study material in addition to cell lines derived from Flp-In 293 T-Rex cells.

Table 1. Tumor samples included in *MED12* exon 1/2 mutation screenings in Studies I-IV.

Tumor type	*n* (study)	Sample type	Samples from	Reference
Tumors of mesenchymal origin				
Uterine Leiomyosarcoma	**40**			
Early onset	27 (I), 24 (II)	FFPE	Finland, PH/CFCH	(Ylisaukko-oja *et al.*, 2006)
Unselected	12 (I), 13 (II)			
Sarcoma	**104**			
Soft tissue sarcoma	83 (I), 79 (II)	FF	Finland, PH	
Bone sarcoma	21 (I), 19 (II)			
Gastrointestinal stromal tumor	12 (I)	FFPE	Finland, CFCH	
Uterine leiomyoma	**353**			
Conventional	73(II),66 (III)	FF / FFPE	Finland, HUCH and PH / South Africa, UCT	(Mäkinen *et al.*, 2011a; Mäkinen *et al.*, 2011b; Mäkinen *et al.*, 2013)
Cellular	49 (II)			
With bizarre nuclei (Atypical)	15 (II)			
Mitotically active	15 (II)			
From HLRCC patients	34 (II),122 (III)			
From South African patients	13 (II)			
Extrauterine leiomyoma	42 (I), 39 (II)	FFPE	Finland, PH/CFCH	(Kiuru *et al.*, 2002)
Endometrial polyp	54 (I), 55 (II)	FFPE	Finland, PH	
Lipoma	35 (I)	FFPE	Finland, CFCH	
Estrogen-progesterone dependent tumors				
Ovarian carcinoma	**122** (I)	FF / FFPE	Finland, PH	
Clear cell	39			
Serous	44			
Mucinous	10			
Endometrioid	10			
NOS	19			
Breast cancer	**94** (I)	FF	Finland, OGH/PH	
Ductal	68			
Lobular	14			
Medullary	4			
Other	8			
Hematological malignancies				
Acute myeloid leukemia	131 (I)	Fresh / FF	Denmark, AAUH / US, MSKCC	
Acute lymphoblastic leukemia (T-cell origin)	37 (I)	FF	Denmark, AAUH	
Myeloproliferative neoplasm	**96** (I)	Fresh / FF	US, MSKCC	
Polycythemia vera	48			
Essential thrombocytosis	48			
Chronic lymphocytic leukemia	**746** (IV)	Fresh / FF / Fixed	US, CRC US, OSU Finland, HUCH	(Rassenti *et al.*, 2004; Rassenti *et al.*, 2008)
	278			
	292			
	176			
Tumors with aberrant Wnt-signaling				
Colorectal cancer	392 (I), 183 (II)	FF	Finland, FCH	(Aaltonen *et al.*, 1998; Salovaara *et al.*, 2000)

FFPE, formalin-fixed paraffin-embedded; FF, frozen while fresh; PH, Department of Pathology at Helsinki University Central Hospital; CFCH, Central Finland Central Hospital; HLRCC, Hereditary Leiomyomatosis and Renal Cell Cancer; HUCH, Helsinki University Central Hospital; UCT, Department of Obstetrics and Gynecology at University of Cape Town; NOS, not otherwise specified; OGH, Department of Obstetrics and Gynecology at Helsinki University Central Hospital; AAUH, Department of Hematology at Aarhus University Hospital; MSKCC, Memorial Sloan-Kettering Cancer Center; CRC, Chronic lymphocytic leukemia Research Consortium; OSU, Ohio State University Sample Bank

Table 2. Finnish HLRCC families included in Study III.

HLRCC family	No. of family members included in the study	No. of tumor samples included in the study	Germline *FH* mutation	Reference
B	3	19	c.671_672delAG; p.E224fs	(Launonen *et al.*, 2001)
C	3	5	c.1027C>T; p.R343X	(Kiuru *et al.*, 2001)
D	2	19	c.587A>G; p.H196R	(Lehtonen *et al.*, 2006)
E	1	6	c.587A>G; p.H196R	(Lehtonen *et al.*, 2006)
M	9	23	c.671_672delAG; p.E224fs	(Launonen *et al.*, 2001)
N	1	6	c.1027C>T; p.R343X	(Koski, 2010)
KH-80	1	5	c.587A>G; p.H196R	
MG_56	2	12	c.671_672delAG; p.E224fs	
OuluM1	3	6	c.583A>G; p.M195V	(Tolvanen *et al.*, 2012)
My5006	1	15	c.583A>C, p.M195V	(Heinonen *et al.*, 2014)
My31	1	6*	c.1439C>G, p.S480X	(Mäkinen *et al.*, 2014b)

HLRCC, Hereditary Leiomyomatosis and Renal Cell Cancer; fs, frameshift; *additional 24 archival tissue specimens

1.2 Normal tissue samples (I-III)

To validate the somatic origin of the mutations observed in the first screenings, the genomic site was also evaluated from the corresponding normal tissue sample of each patient. Normal tissue samples were analyzed from two patients with uterine leiomyoma and two colorectal cancer patients in Study I and from five patients with uterine leiomyoma in Study II. Corresponding normal myometrium specimens of the tumor samples included in the gene expression analyses were utilized in Studies II and III.

1.3 Cell lines (II, V)

A cell line originating from human embryonic kidney cells, HEK293T (American Type Culture Collection [ATCC], Manassas, VA, USA), was used to express MED12 wild-type (WT) and mutant derivatives transiently in Study II. Flp-In 293 T-Rex cell line (Invitrogen, Life technologies, Carlsbad, CA, USA) containing a flippase recognition target site at a transcriptionally active locus was utilized to create stable and inducible MED12 WT and mutant expressing cell lines in Study V. The cell line was authenticated with the Promega GenePrint10 System (Promega, Madison, WI,

USA) at the Genomics Unit of Technology Centre, Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland.

2. Methods

2.1 Histopathological evaluation (I, III)

Histopathological evaluation of the tumor samples was performed by the collaborating pathologists to confirm the diagnosis (Study I) and an adequate tumor percentage of the samples, and to select representative areas of the tumor tissue for tissue microarray (TMA; Study III). For the evaluation, 5-µm sections were cut from the FFPE blocks and stained with hematoxylin and eosin (HE; Table 3).

In Study I, pathologist Dr. Johanna Arola evaluated and classified previously collected series of early-onset uterine leiomyosarcoma samples according to the new criteria defined by the World Health Organization (WHO) (Tavassoli and Devilee, 2003; Ylisaukko-oja *et al.*, 2006). Additional unselected uterine leiomyosarcoma samples and soft tissue sarcoma series including two uterine leiomyosarcoma metastases were evaluated by pathologists Dr. Jan Böhm and Prof. Tom Böhling, respectively.

In Study III, all available uterine tumor specimens from the hysterectomies performed on 27 Finnish HLRCC patients and 66 sporadic uterine leiomyoma samples were evaluated, and representative tumor areas marked by gynecological pathologist Dr. Ralf Bützow.

2.2 DNA and RNA extraction (I-V)

Genomic DNA was extracted from fresh, fresh frozen, or fixed tissue samples utilizing either the conventional non-enzymatic method (Miller *et al.*, 1988; Lahiri and Nurnberger, 1991) (I,II,IV), the standard phenol-chloroform method (IV), or the commercial extraction kits FastDNA® (MP Biomedicals LLC, Solon, OH, USA; I-III), DNeasy® Blood and Tissue (Qiagen, Hilden, Germany; I,IV,V), and MagNA Pure LC DNA Isolation Kit I (Roche, Basel, Switzerland; I). Proteinase K digestion followed by phenol-chloroform isolation or the NucleoSpin® DNA FFPE XS Kit (Macherey-Nagel GmbH & Co KG, Düren, Germany) were utilized to extract genomic DNA from FFPE tissue samples (I-III).

Total RNA was extracted from fresh frozen samples using TRIzol® Reagent (Invitrogen) or TRI Reagent® (Molecular Research Center Incorporated, Cincinnati, OH) (II, III). Extracted RNA was purified with RNeasy® MinElute™ Clean Up Kit (Qiagen) for gene expression analysis. cDNA synthesis was performed with standard

methods using Moloney Murine Leukemia Virus (M-MLV) reverse transcriptase enzyme and random primers (Promega).

2.3 Sanger sequencing (I-V)

Oligonucleotide primers used in polymerase chain reaction (PCR) were designed with Primer3 or Primer3Plus (Untergasser *et al.*, 2012). PCR was performed according to a standard protocol, and the products were purified using ExoSAP-IT (USB Corporation, Cleveland, OH) or A'SAP (ArcticZymes, Tromsø, Norway). Sequencing was performed on an Applied Biosystems ABI3730 Automatic DNA Sequencer (Life technologies, Thermo Fisher Scientific, Waltham, MA, USA) at the Technology Center, Institute for Molecular Medicine Finland (FIMM), Helsinki, Finland. Sequences were analyzed manually using FinchTV and with the Mutation Surveyor software (Softgenetics, State College, PA, USA).

2.3.1 Mutation screening and validation (I-V)

MED12 exon 2 mutation screening in Study I was performed with direct Sanger sequencing. In Study II, the samples were similarly screened for *MED12* exon 1 mutations. Sanger sequencing of cDNA was applied to four mutation-positive uterine leiomyoma samples from which tumor tissue was available. Somatic status of the mutations observed in Studies I and II was verified by sequencing the corresponding normal tissue samples. Mutation screenings performed in Studies III and IV included sequencing of both *MED12* exon 1 and 2. The sample series included fresh/fresh frozen/fixed and FFPE samples from which 15 and 25 ng of genomic DNA, respectively, was used in PCR. In Study V, both genomic DNA and cDNA from the T-ALL patient sample were sequenced to validate the presence and expression of the *MED12* c.97G>T, p.E33X mutation.

2.3.2 Loss of heterozygosity analysis (III)

LOH at the *FH* locus was analyzed from tumor DNA (fresh frozen samples) of the HLRCC patient with multiple *MED12* mutation-positive tumors. Five independent PCR reactions per sample were sequenced and the heights of the wild-type and mutant peaks in the chromatograms were compared manually. LOH was recorded when the height of the wild-type allele peak was repeatedly reduced when compared to the height of the mutant allele peak.

2.4 Gene expression analysis (II, III)

Gene expression analysis to examine the global expression profiles and clustering of the tumors was performed utilizing Affymetrix GeneChip Human Exon 1.0 ST Arrays (Affymetrix, Santa Clara, CA) at the Biomedicum Functional Genomics Unit (FuGU), Helsinki, Finland. Re-mapped Brainarray Custom CDF files (HuEx10stv2_Hs_ENSG, Versions 16 and 17) were used when analyzing data with Partek Genomic Suite™ v. 6.5 (Partek Incorporated, St. Louis, MO, USA). All samples were quantile-normalized by the Robust Multichip Average (RMA) method and adjusted for probe sequence and GC-content. Clustering of the tumors was evaluated with unsupervised hierarchical clustering analysis (Cosine dissimilarity), performed with the most variable 1% of genes (n=372), defined by the coefficient of variation calculated across all tumor samples.

2.5 Immunoprecipitation and kinase activity assay (II)

Immunoprecipitation and kinase activity assays were performed in Prof. Thomas Boyer's lab at the University of Texas, Healt Science Center at San Antonio, Texas, USA. *MED12* exon 1 mutations were created with site-directed mutagenesis (SDM) to *MED12* cDNA (coding for amino acids 1-593) containing pCDNA3.1 plasmids and further cloned to p3xFLAG-MED12 expression vectors. Plasmids were transfected to HEK293T cells using X-tremeGENE 9 transfection reagent (Roche), and after 48 h MED12 WT and mutant derivatives were immunoprecipitated from cell lysates using anti-FLAG M2 affinity gel (Sigma-Aldrich, St. Louis, MO). Western blot analysis of the immunoprecipitates was performed with antibodies designated in Table 3. [ɣ-^{32}P]-ATP and purified glutathione S-transferase (GST) RNA pol II CTD recombinant peptide were used to detect CTD-directed kinase activity of the immunoprecipitates. Exon 2 mutation L36R was included in the analysis as a control.

2.6 Western blotting (II, V)

Extracted proteins were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) using 10% Tris-HCl polyacrylamide gels and transferred to polyvinylidene difluoride (PVDF) membranes (Bio-Rad Laboratories, Hercules, CA, US). The antibodies used in protein detection are listed in Table 3. Antibody binding was visualized using Pierce™ ECL Western Blotting Substrate (Thermo Fisher Scientific; II) or ChemiDoc™ MP System (Bio-Rad Laboratories; V) and quantified with ImageQuant software (GE Healthcare, Litte Chalfont, UK).

2.7 Tissue microarray construction and immunohistochemistry (III)

A manual tissue arrayer (MTA-I, Beecher Instruments, Sun Prairie, WI, USA) was used to construct TMAs containing 83 uterine leiomyomas from HLRCC patients and 66 sporadic uterine leiomyomas. Four 0.8-mm tumor cores were punched from the representative tumor area in the original sample block and inserted into a newly casted paraffin block. Cores representing normal myometrium tissue were included in the TMAs to serve as internal controls.

Immunohistochemical analysis (IHC) of *FH* inactivation was performed on 5-µm thick tissue sections of TMA or FFPE or Optimal Cutting Temperature compound (OCT) embedded tissue samples. The EnVision™ + kit (Dako, Agilent Technologies, Santa Clara, CA) with anti-2SC antibody and anti-rabbit horseradish peroxidase (HRP) polymer (Table 3) was used in the analysis. The samples were scored positive when displaying both nuclear and cytoplasmic staining and negative when no staining or only low cytoplasmic staining of individual cells was observed.

2.8 Creating MED12-expressing Flp-In 293 T-Rex cell lines (V)

Expression vectors with *MED12* mutations were created from a destination vector (pTO_HA_StrepIII_c_GW_FRT) containing *MED12* cDNA (Varjosalo *et al.*, 2013) using the QuickChange Site-Directed Mutagenesis kit (Agilent Technologies). For the BioID method, mutations were created in *MED12* cDNA in Gateway pDONR221 entry clones and further cloned into pTO_MYC_BirA_C vectors. Flp-In 293 T-Rex cells were co-transfected with the construct containing expression vectors and the Flp recombinase expression vector (pOG44; Invitrogen, Life Technologies) using FuGENE® HD Transfection Reagent (Promega). Selection with Hygromycin B (Invitrogen, Life Technologies) to create stable construct-expressing cell lines was started two days after the transfection. Stable cell lines were cultured in Dulbecco's modified Eagle medium (Biowhittaker® DMEM 4.5 g/L glucose; Lonza, Basel, Switzerland) supplemented with 10% fetal bovine serum, 2 mM L-alanyl-L-glutamine, 50 mg/ml penicillin, 50 mg/ml streptomycin, 15 µg/ml blasticidin, and 100 µg/ml Hygromycin B.

2.9 Immunofluorescence (V)

Flp-In 293 T-Rex cells expressing HA-tagged (tag derived from human influenza hemagglutinin) WT or mutant MED12 derivatives were analyzed with immunofluorescence (IF) staining. The cells were plated on coverslips treated with Poly-L-lysine hydrobromide (Sigma-Aldrich) and induced to express the *MED12* construct with tetracycline (1 µg/ml Doxycycline, Sigma-Aldrich). After 24 h culture, cells were fixed with 4% paraformaldehyde and stained using the antibodies and dyes indicated in Table 3. The Axioplan 2 upright epifluorescence microscope and

LSM 780 confocal microscope with Plan-neofluar 40x, 1.3 NA oil objective (Carl Zeiss, Oberkochen, Germany) were used to visualize and image the cells at the Biomedicum Imaging Unit, University of Helsinki, Helsinki, Finland. Two serial optical sections were combined and used for image display. Brightness and contrast was adjusted and the images were compiled with Photoshop CS5.1 (Adobe, San Jose, CA).

The immunofluorescence images were analyzed with an Anima platform-based pipeline (Rantanen *et al.*, 2014). DAPI (4', 6-diamidino-2-phenylindole) staining was used to determine the areas of the cell nuclei, and the cytoplasm was defined as a 20-pixel wide ring around each nucleus. Signal intensities of HA-tagged MED12 staining within the nuclei and the cytoplasms were measured, and the ratio of these intensities (relative nucleus/cytoplasm margination) was calculated.

Table 3. Antibodies and dyes used in studies II, III, and V

Antibody	Manufacturer	Method used in	Study
Primary antibodies			
Anti-FLAG M2	F3165; Sigma-Aldrich	WB	II
Anti-MED23	550429; BD Biosciences	WB	II
Anti-CDK19	HPA007053; Sigma-Aldrich	WB	II
Anti-CDK8	sc-1521; Santa Cruz Biotechnology	WB	II
Anti-MED4	Kim et al. (Kim *et al.*, 2006)	WB	II
Anti-Cyclin C	558903; BD Biosciences	WB	II
Anti-2SC	Bardella et al. (Bardella *et al.*, 2011)	IHC	III
Anti-HA.11	16B12; Covance	WB, IF	V
Anti-MED12	sc-5374, Santa Cruz Biotechnology	WB	V
Anti-Vinculin	sc-5573, Santa Cruz Biotechnology	WB	V
Secondary antibodies			
Anti-mouse IgG-Peroxidase	A4416; Sigma–Aldrich	WB	II,V
Anti-goat IgG-Peroxidase	A5420; Sigma–Aldrich	WB	II,V
Anti-rabbit IgG-Peroxidase	A6154; Sigma–Aldrich	WB	II,V
Anti-rabbit Ig-HRP polymer	Dako, Agilent Technologies	IHC	III
Anti-mouse IgG-Alexa-488	R37120; Thermo Fisher Scientific	IF	V
Dyes			
Hematoxylin		HE	I,III
Eosin		HE	I,III
DAB+ (3,3'-diaminobenzidine chromogen)	Dako, Agilent Technologies	IHC	III
F-actin probe (Phalloidin) conjugated Alexa Fluor®568	Thermo Fisher Scientific	IF	V
DAPI (4', 6-diamidino-2-phenylindole)	DAPI Vectashield® Mounting Media, Vector Laboratories	IF	V

WB, Western blotting; IHC, immunohistochemistry: IF, immunofluorescence; HE, hematoxylin and eosin staining; Covance, Princeton, NJ, USA; Santa Cruz Biotechnology, Santa Cruz, CA; BD Biosciences, San Diego, CA; Vector Laboratories, Burlingame, CA

2.10 Affinity purification and BioID -mass spectrometry (V)

Proteome-wide protein-protein interactions were analyzed from Flp-In 293 T-Rex cells expressing WT or mutant MED12 derivatives with liquid chromatography-mass spectrometry (LC-MS/MS) preceded by single-step affinity purification (AP). The BioID method, based on biotinylation of the nearby proteins and subsequent mass spectrometry analysis, was also utilized to detect more transient interactions (Roux *et al.*, 2012). Analyses were performed in collaboration with docent Markku Varjosalo at the Institute of Biotechnology, University of Helsinki, Helsinki, Finland. Three biological replicates, each consisting of tetracycline-induced cells from 5 x 15-cm plates, were processed for each MED12 derivative-expressing cell line. For cells analyzed with BioID, an additional 50 μM biotin was added 24 h prior the collection. A detailed description of the AP and LC-MS/MS methods can be found in the group's recent publications (Varjosalo *et al.*, 2013; Turunen *et al.*, 2014). In brief, cells were lysed with HNN lysis buffer and additionally in the BioID method by sonication. The cleared lysates were loaded into spin columns (Bio-Rad Laboratories) with Strep-Tactin beads (IBA GmbH, Göttingen, Germany) to isolate the Strep-tagged bait proteins and bound interactors. The beads were washed with HNN lysis buffer and with plain HNN buffer, after which the proteins were eluted with D-biotin (Thermo Fisher Scientific) in HNN buffer. For the LC-MS, samples were pretreated to reduce and alkylate the cysteine bonds and digested overnight at 37 °C with trypsin (Promega). Samples were further purified and re-dissolved in suitable buffers. The analyses were performed on an EASY-nLCII-system coupled to Orbitrap Elite ETD hybrid mass spectrometer using the Xcalibur version 2.2 SP 1.48 (all from Thermo Fisher Scientific) via a nanoelectrospray ion source as described in Varjosalo *et al*., 2013. Peak extraction and subsequent protein identification against the human reference proteome of the UniProtKB/SwissProt database was performed utilizing the SEQUEST search algorithm in Proteome Discoverer software (Thermo Fisher Scientific). Filtering against the control contaminant database was used to identify high confidence protein-protein interactions. Relative protein abundances (normalized to the bait) were calculated from the spectral counts (average and standard deviation of three replicates).

2.11 *In silico* prediction tools and online databases (I, IV, V)

Functional effects of the non-synonymous changes observed in Study I were evaluated using the online *in silico* prediction tools PolyPhen2 and SIFT. Multiple Sequence Comparison by Log-Expectation (MUSCLE) was applied as a sequence alignment tool to evaluate the evolutionary conservation of the MED12 amino (N)-terminal region in Studies I and V. COSMIC database was utilized in Studies IV and V to perform a systematic search for *MED12* mutations in several different cancer types and to analyze the presence of frame-altering or nonsense mutations in the gene. In Study V, four *in silico* prediction tools were used to evaluate the presence of a nuclear localization signal (NLS) on the N-terminal region of MED12 (amino acids 1-154,

indicated to be lost through c.97G>T, p.E33X mutation). The UniProtKB/SwissProt database was utilized in peptide/protein identification after MS analyses. All *in silico* prediction tools and online databases used in this study are listed in Table 4.

Table 4. *In silico* prediction tools and online databases utilized in studies I, IV, and V

Application	Database / Software	Reference	Study
Amino acid substitution effect prediction	PolyPhen2	(Adzhubei *et al.*, 2010)	I
	SIFT	(Kumar *et al.*, 2009)	I
Peptide sequence alignment	MUSCLE	(Edgar, 2004)	I,V
Mutation database	COSMIC	(Forbes *et al.*, 2008)	IV,V
NLS prediction	SeqNLS	(Lin and Hu, 2013)	V
	PSORT II	(Nakai and Horton, 1999)	V
	cNLS Mapper	(Kosugi *et al.*, 2009)	V
	NLStradamus	(Nguyen Ba *et al.*, 2009)	V
Protein identification	UniProtKB/SwissProt	(UniProt Consortium, 2015)	V

2.12 Statistical analyses (II-V)

Statistical analyses were performed with R software, version 3.0.2 (R Foundation for Statistical Computing, Vienna, Austria), and with Python, version 2.7 (Python Software Foundation, Wilmington, DE, USA). In Study II, Student's t-test was utilized to evaluate the statistical significance of the differences in the levels of ^{32}P-GST-CTD between the MED12 WT and mutant derivative immunoprecipitates observed in the kinase activity assay.

Fisher's exact test (two-sided *p*-value) was used in Study III to test the statistical significance of the differences in the frequencies of *MED12* mutations and biallelic *FH* inactivation between uterine leiomyomas from HLRCC patients and sporadic conventional uterine leiomyomas. A permutation test was utilized to evaluate the mutual exclusiveness of *MED12* mutations and FH deficiency in uterine leiomyomas. Observed mutations were randomly redistributed between analyzed samples in 1,000,000 permutations. An empirical *p*-value was computed as $p = (1+k)/n$, where k is the number of permutations in which at least one sample was observed with both changes and n is the number of permutations. To evaluate the statistical significance of many *MED12* mutation-positive tumors co-occurring in HLRCC patient My31, tumors were redistributed to patients in 10,000,000 permutations. Here, k denoted the number of permutations in which six or more *MED12* mutation-positive tumors and no FH-deficient tumors were distributed to the same patient.

In Study IV, Fisher's exact test was applied to evaluate the statistical significance of the difference observed in the accumulation of the *MED12* mutations to exons 1 and 2 between CLL and uterine leiomyomas. Associations of the *MED12* mutation status with age, sex, the mutation status of immunoglobulin heavy chain variable (IGHV) genes, CD38 expression, 70 kD zeta-associated protein (ZAP-70) expression, and

methylation status of ZAP-70 were analyzed using Fisher's exact test (categorical variables) or Wilcoxon rank sum test (continuous variables).

Statistical significance of the differences in the nucleus/cytoplasm signal intensity ratios between cells expressing MED12 WT and mutant derivatives (G44D, E33K, and E33X) were evaluated using the Wilcoxon rank sum test in Study V. Student's t-test was used to evaluate the statistical significance of the differences seen in the abundances of bound nuclear pore proteins between the MED12 WT and mutant derivatives (E33X, NLS1, and NLS2) in the BioID analysis.

3. Ethical issues

All samples included in the study were collected either after signed informed consent or after the approval by the National Supervisory Authority for Welfare and Health (Finland). Anonymized samples used in Studies I-III were collected after the approval by the director of the health care unit. Authorization to use samples of deceased Finnish CLL patients was obtained from the Ethics Committee of the Helsinki University Central Hospital. All studies were approved by the appropriate Ethics Committees of the Hospital District of Helsinki and Uusimaa, Finland, as well as respective local ethics committees.

RESULTS

1. *MED12* exon 2 mutations in uterine leiomyosarcoma and colorectal cancer (I)

To evaluate the role of *MED12* exon 2 mutations in various tumor types other than uterine leiomyomas, Sanger sequencing-based mutation screening was performed on 1158 tumor samples representing 14 different tumor types, both benign and malignant (see Table 1 for detailed information on the sample set). As a result, five *MED12* exon 2 mutations were identified, three in uterine leiomyosarcomas and two in colorectal cancers.

1.1 *MED12* exon 2 mutations occur recurrently in uterine leiomyosarcoma

MED12 exon 2 mutations were observed in uterine leiomyosarcomas with a frequency of 7%, as a mutation was identified in three out of 41 samples analyzed (Table 5, Figure 4). Two of the mutations were observed in samples representing early onset (age at diagnosis ≤ 45 years) uterine leiomyosarcomas. The third was identified in a sample originally screened as a soft-tissue sarcoma, but further diagnosed as a uterine leiomyosarcoma metastasis. All three mutations occurred in the same, highly conserved region at the beginning of exon 2, where effectively all *MED12* mutations are located in uterine leiomyomas. One of the mutations affected codon 44, the most commonly mutated codon in uterine leiomyomas, leading to a substitution of the glycine amino acid to serine (c.130G>A, p.G44S). Two other mutations were an insertion of 21 nucleotides (c.115_116ins21, p. A38_L39ins7) and a deletion of three nucleotides (c.104_106delAAC, p. E35_L36delinsV). Both of these changes represented frame-retaining insertion/deletion mutations typically seen in uterine leiomyomas. The somatic origin of the mutations observed in early-onset uterine leiomyosarcomas were verified by sequencing the DNA extracted from the corresponding normal tissue samples.

1.2 Rare *MED12* exon 2 mutations in colorectal cancer

In total, 392 colorectal cancer samples were screened for *MED12* exon 2 mutations. Only two mutations were identified (2/392; 0.5%) as a result of the screening (Table 5, Figure 4). One of these mutations affected the mutational hotspot codon 44. This mutation (c.130G>T, p.G44C) was identified in a sample from a female patient with a Dukes B/grade II/microsatellite stable tumor. Another missense change (c.200C>T, p.A67V) was observed in a sample from a female patient with a Dukes A/grade II tumor displaying microsatellite instability. Amino acid alanine 67 is situated at an evolutionary highly conserved region at the MED12 N-terminus, but no mutations

affecting this particular amino acid had been reported prior to this study. The significance of the variant remains uncertain also based on the *in silico* prediction tools PolyPhen2 and SIFT, both of which estimate A67V substitution to be tolerated. Corresponding normal tissue samples were sequenced, and both mutations observed in these colorectal cancer cases were verified as somatic.

Table 5. *MED12* exon 2 mutations observed in the screening of 1158 tumor samples.

Tumor type	Sample ID	*MED12* ex 2 mutation
Uterine leiomyosarcoma	LM_62.1T	c.130G>A, p.G44S
Uterine leiomyosarcoma	LM_53.1T	c.115_116ins21, p. A38_L39insSHDELTA
Uterine leiomyosarcoma (metastasis)	GE05-129	c.104_106delAAC, p.E35_L36delinsV
Colorectal cancer	c834.1T	c.130G>T, p.G44C
Colorectal cancer	c43.1T	c.200C>T, p.A67V

Figure 4. Sequence chromatograms of *MED12* exon 2 mutations observed in three uterine leiomyosarcoma (A) and two colorectal cancer (B) samples. The mutation sites are indicated by arrows in the chromatograms of the tumor samples. Sequence chromatograms of the corresponding normal tissues are shown in the lower lanes (not available for the GE05-129 sample).

2. Mutations in exon 1 of *MED12* (II)

After the identification of highly frequent *MED12* mutations in uterine leiomyomas, subsequent Sanger sequencing-based mutation screenings had concentrated on exon 2 and the preceding intron-exon boundary. To analyze whether mutations occur also in the first exon of the gene, 611 tumor samples were screened for *MED12* exon 1 mutations. The sample set represented tumor types where *MED12* exon 2 mutations have previously been identified (uterine leiomyomas [conventional, various histopathological subtypes, from HLRCC patients], extrauterine leiomyomas, endometrial polyps, uterine leiomyosarcomas, soft-tissue and bone sarcomas, and colorectal cancers; see Table 1 for detailed information on the sample set). All individual samples included in the screening were confirmed as *MED12* exon 2 mutation-negative.

2.1 *MED12* exon 1 mutations in conventional uterine leiomyomas

Five mutations located in exon 1 of *MED12* were identified as a result of the screening (5/611; 0.8%). All observed mutations were in-frame insertion/deletion mutations in conventional uterine leiomyomas (5/86; 5.8%; Table 6). Four of the mutations were observed in samples from Caucasian patients and one in a uterine leiomyoma from a South African patient with mixed ancestry (colored) (Mäkinen *et al.*, 2011a). The somatic status of all five mutations was verified through normal tissue DNA sequencing. Frozen specimens of the tumor tissue were available from four mutation-positive tumors of Caucasian patients. cDNA sequencing showed predominant expression of the mutant allele in each tumor (Figure 1A in the original publication II). No mutations were observed in any other tumor type included in the screening.

Table 6. Five *MED12* exon 1 mutations observed in conventional uterine leiomyomas.

Ethnicity of the patient	Sample ID	*MED12* ex 1 mutation
Caucasian	M4m2	c.76_91del16insG, p.P26_Q31delInsE
Caucasian	MY24m6	c.82_99del18, p.D28_E33del
Caucasian	MY32m13	c.82_99del18, p.D28_E33del
Caucasian	MY33m4	c.80A>T; c.84_98del15, p. Q27L; p.D28_K32del
South African, Mixed origin	FG166_1	c.77C>T; c.79_99del21, p.P26L; p.Q27_E33del

2.2 *MED12* exon 1 mutations lead to similar gene expression profile as exon 2 mutations

To evaluate the functional impact of *MED12* exon 1 mutations, gene expression analysis and immunoprecipitation accompanied with kinase activity assays were performed. Gene expression profiling was applied to four exon 1 mutation-positive tumors where RNA from the tumor tissue and corresponding normal myometrium

tissue was available. Data from these *MED12* exon 1 mutation-positive tumors were analyzed together with the data from previous study comprising 16 *MED12* exon 2 mutation-positive tumors, four FH-deficient tumors, 10 *HMGA1/HMGA2* overexpressing tumors, and eight tumors without any of these aberrations (Mehine *et al.*, 2013). All *MED12* mutation-positive tumors displayed similar patterns of gene expression, with *RAD51B* being the most upregulated gene. The exact location of the mutation at the *MED12* 5' end did not affect the clustering of the tumors as all *MED12* mutation-positive tumors clustered clearly together in unsupervised hierarchical clustering (Figure 5).

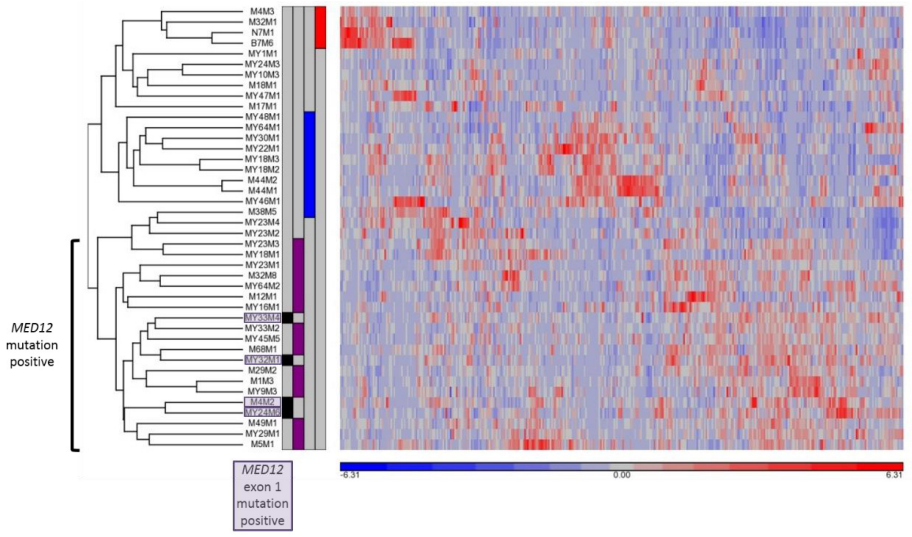

Figure 5. Unsupervised hierarchical clustering analysis of 42 leiomyomas from 31 patients. Based on hierarchical clustering analysis, four *MED12* exon 1 mutation-positive leiomyomas (black) clustered together with 16 *MED12* exon 2 mutation-positive leiomyomas (purple). The data were analyzed together with data from four FH-deficient leiomyomas (red), 10 *HMGA1/HMGA2* overexpressing leiomyomas (blue), and eight leiomyomas missing all these aberrations.

2.3 MED12 exon 1 mutations disrupt the Mediator kinase module integrity

Possible effects of the *MED12* exon 1 mutations on the protein-protein interactions within Mediator were analyzed utilizing immunoprecipitation. Similarly as seen previously with exon 2 mutations (Turunen *et al.*, 2014), interactions between Mediator kinase module components were severely disrupted. MED12 constructs containing exon 1 mutations (and an additional exon 2 mutant construct as a control), did not co-immunoprecipitate kinase module partners CDK8/19 nor Cyclin C as MED12 WT did (Figure 6A). Two subunits of the Mediator core, MED4 and MED23, were co-immunoprecipitated with all the constructs, indicating that the interaction of MED12 with the Mediator core remained intact despite the mutations in exons 1 and 2 (Figure 6A). Kinase activity assays measuring the RNA-polymerase II CTD directed activity showed significantly diminished levels of activity with mutant

immunoprecipitates compared to that of the MED12 WT (Figure 6B). Again, an identical effect was observed with constructs containing *MED12* exon 1 mutations as was seen with the exon 2 mutation-positive control and in the previous study (Turunen *et al.*, 2014).

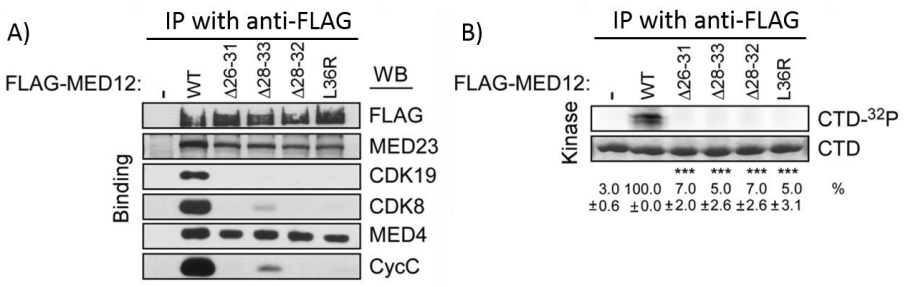

Figure 6. Immunoprecipitation followed by Western blot analysis (A) and *in vitro* kinase assay (B). MED12 derivatives harboring an exon 1 or 2 mutation loses interactions with Mediator kinase module components and hence the RNA pol II CTD directed kinase activity of the module. CTD-^{32}P levels are expressed relative to the level in the MED12 WT immunoprcipitate (average ± SEM of three independent experiments). Asterisks denote statistically significant differences versus wild type, $***p < 0.001$.

3. The role of *MED12* in HLRCC patients' uterine leiomyomas (III)

In Study III, the aim was to evaluate the contribution of *MED12* mutations to the tumorigenesis of HLRCC patients' uterine leiomyomas and to analyze whether *MED12* mutations and the most prevalent tumorigenic mechanism in HLRCC-associated tumors, biallelic *FH* inactivation, are mutually exclusive.

3.1 *MED12* mutations and FH deficiency are mutually exclusive in uterine leiomyomas

2SC immunohistochemistry, indirectly indicating FH deficiency, confirmed that the great majority of HLRCC patients' uterine leiomyomas (113/122 uterine leiomyomas from 26 individuals; 92.6%) display biallelic inactivation of *FH*. Only nine tumors (from four different patients) showed negative 2SC staining. Interestingly, *MED12* exon 1 and 2 mutation screening revealed exon 2 mutations in all of these 2SC-negative tumors (9/122; 7.4%; Table 7 and Figures 7 and 8). In the set of 66 sporadic uterine leiomyomas, 35/64 (55%) successfully sequenced tumors harbored *MED12* exon 1/2 mutation (Figure 8). All sporadic tumors were scored negative in 2SC immunohistochemical analysis, and were thus proficient for FH activity. The difference in the observed frequencies of *MED12* mutations and biallelic *FH* inactivation between HLRCC patients' and sporadic uterine leiomyomas was highly significant ($p < 2.2 \times 10^{-16}$). Statistical analysis of the data using a permutation test

(1,000,000 permutations) also validated *MED12* mutations and biallelic *FH* inactivation to be mutually exclusive ($p = 1x10^{-6}$).

Table 7. Four HLRCC patients with at least one *MED12* mutation-positive tumor.

Patient	Germline *FH* mutation	Tumors included in the screening *MED12* mutation status	2SC IHC*
B3	c.671_672delAG, p.E224fs	wt	+
		c.131G>A, p.G44D	-
		wt	+
B7	c.671_672delAG, p.E224fs	c.113-160del48, p.A38_G53del	-
		wt	+
		wt	+
		wt	+
		wt	+
		wt	+
E1	c.587A>G, p.H196R	wt	+
		wt	+
		wt	+
		wt	+
		wt	+
		c.130G>A, p.G44S	-
My31	c.1439C>G, p. S480X	c.130G>A, p.G44S	-
		c.131G>T, p.G44V	-
		IVS1-1_139del41	-
		IVS1-8T>A, p.E33_D34insPQ	-
		c.131G>A, p.G44D	-
		c.130G>C, p.G44R	-

fs, frameshift; *2SC IHC + = FH-deficient tumor; 2SC IHC - = FH-proficient tumor

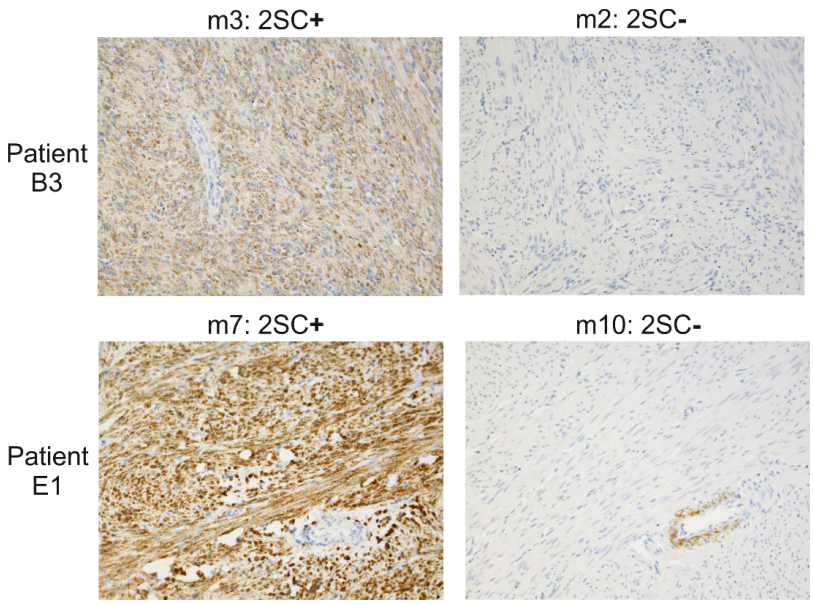

Figure 7. 2SC IHC staining of four uterine leiomyomas from two HLRCC patients (original magnification, x 200). FH-deficient, *MED12* mutation-negative tumors display positive 2SC staining (left panel), whereas *MED12* mutation-positive tumors are FH proficient and show negative 2SC staining (right panel).

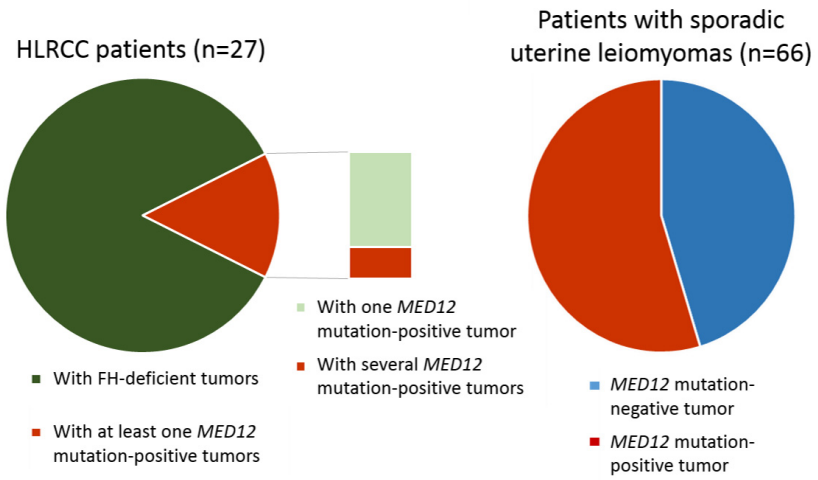

Figure 8. Occurrence of FH-deficient, *MED12* mutation-positive, and *MED12* mutation-negative tumors in HLRCC patients and in patients with sporadic uterine leiomyomas. The majority of the HLRCC patients harbored only FH-deficient tumors. Out of the four patients with at least one *MED12* mutation-positive tumor, only one had several *MED12* mutation-positive tumors. Patients with sporadic tumors harbored tumors with or without *MED12* mutation, and no FH-deficient tumors were observed.

3.2 *MED12* mutation-positive uterine leiomyomas from HLRCC patients display similar gene expression profile as sporadic *MED12* mutation-positive tumors

Gene expression analysis was performed to compare the expression profiles of HLRCC patients' *MED12* mutation-positive tumors with those of FH-deficient tumors and sporadic *MED12* mutation-positive tumors. Data from 26 *MED12* mutation-positive tumors (seven from HLRCC patients), 10 FH-deficient, and 39 *MED12* and *FH* WT tumors as well as their corresponding normal myometrial tissue samples were analyzed together (Mehine *et al.*, 2013; Mehine *et al.*, 2016). Regardless of the germline *FH* status, all tumors harboring a *MED12* mutation clustered together in unsupervised hierarchical clustering (Figure 3 in the original publication III). FH-deficient tumors (both from HLRCC patients and sporadic tumors with two somatic *FH* inactivating aberrations) formed their own separate cluster. Germline *FH* mutation did not affect the expression profile of the normal myometrium samples, as they all clustered together.

3.3 HLRCC patient with multiple *MED12* mutation-positive uterine leiomyomas

MED12 exon 1 and 2 mutation screening combined with 2SC immunohistochemistry pointed out an interesting case where all tumors from one HLRCC patient (My31) were observed to harbor a somatic *MED12* mutation and to be proficient in regard to FH. A different *MED12* mutation was identified in all six tumors included in the original sample set, verifying them as separate lesions. Further analysis of 24 archival uterine leiomyoma tissue specimens of the patient showed all except one to be FH proficient and *MED12* mutation-positive. Only one uterine leiomyoma sample displayed 2SC positivity, indicating biallelic *FH* inactivation, and did not harbor a *MED12* mutation. Based on the screening results, the occurrence of this many *MED12* mutation-positive uterine leiomyomas in one HLRCC patient is highly unlikely ($p = 2x10^{-7}$, permutation test with 10,000,000 permutations).

4. *MED12* mutations in chronic lymphocytic leukemia (IV)

A systematic database search utilizing COSMIC revealed five *MED12* exon 1 or 2 mutations in CLL. Four of these mutations occurred in exon 2, and three of them affected the mutational hotspots observed in uterine leiomyomas (L36 and G44). The fourth exon 2 mutation led to an A59P substitution. One mutation located in exon 1 and affected the last codon of the exon coding for E33. To study the occurrence of *MED12* mutations in CLL more thoroughly, in total 746 samples from the US and Finland were collected and sequenced for *MED12* exon 1 and 2 mutations.

4.1 Somatic *MED12* mutations are recurrent in CLL

Mutation screening was performed successfully in 709 samples, and *MED12* exon 1 and 2 mutations were identified in each of the individual sample series comprising altogether 39 mutations in 37 samples (5.2%; Table 8). As previously, all changes were missense or insertion/deletion mutations retaining the reading frame. A novel mutation hotspot was identified as the last amino acid of exon 1, glutamic acid (E) 33, was the second most commonly mutated site after G44 in exon 2. Of the observed mutations, 27/39 (69%) situated in exon 2 and the rest (12/39; 31%) occurred in exon 1 of the gene. The distribution between exon 1 and 2 mutations in CLL (results of this screening combined with the information from COSMIC) is statistically significantly different from the distribution observed in uterine leiomyomas (Mäkinen *et al.*, 2011a; Mäkinen *et al.*, 2011b) ($p = 6.0\text{x}10^{-7}$).

Table 8. CLL samples with *MED12* exon 1 and 2 mutations.

Sample series	*n**	*MED12* mutations exon 1 / exon2	Frequency of mutation-positive samples
In total	**746 / 709**	**12/25[a]**	**5.2%**
CRC (US)	278 / 260	6 / 14[a]	7.7%
OSU Sample bank (US)	292 / 274	2 / 8	3.6%
HUCH (Finland)	176 / 175	4 / 3	4.0%

CRC, Chronic lymphocytic leukemia Research Consortium; OSU, Ohio State University; HUCH, Helsinki University Central Hospital; *number of collected / successfully analyzed samples; [a]two samples harbored two mutations, both in exon 2

4.2 Positive *MED12* mutation status is associated with poor prognosis markers in CLL

Associations between *MED12* mutation status and various molecular and clinical characteristics were analyzed in a CRC sample series (n=260), from which the data were available. A statistically significant association was observed between positive *MED12* mutation status and unmutated IGHV status ($p = 6.8\text{x}10^{-4}$), positive ZAP-70 expression ($p = 7.1\text{x}10^{-4}$), and unmethylated status of ZAP-70 ($p = 5.5\text{x}10^{-4}$), all of which are well-established markers of poor prognosis in CLL. Despite strong associations, effects on overall survival or treatment-free time by the mutation status were not detected. Mutation status did not correlate with the age at diagnosis nor the gender of the patient.

5. Somatic *MED12* nonsense mutation in T-cell acute lymphoblastic leukemia (V)

A *MED12* nonsense mutation, c.97G>T, affecting the last codon of exon 1, E33, was identified in a patient with T-ALL through exome sequencing (Kontro *et al.*, 2014). The mutation was validated by Sanger sequencing of the genomic DNA, and the

expression of the mutant allele was verified from the cDNA of the female patient (Figure 9). This mutation presumably introduces a cryptic splice site, as an additional low-level transcript without the last four nucleotides of exon 1 (r.96_99del) was observed in cDNA sequencing. To evaluate the functional effects of this atypical nonsense mutation, none reported in the 5' end of *MED12* previously, the translated protein product, its localization, and protein-protein interactions were analyzed with several immunochemistry and mass spectrometry-based methods. Stable and inducible Flp-In 293 T-Rex cell lines expressing MED12 WT and E33X mutant as well as E33K (CLL hotspot) and G44D (uterine leiomyoma hotspot) mutant derivatives were created for the analyses.

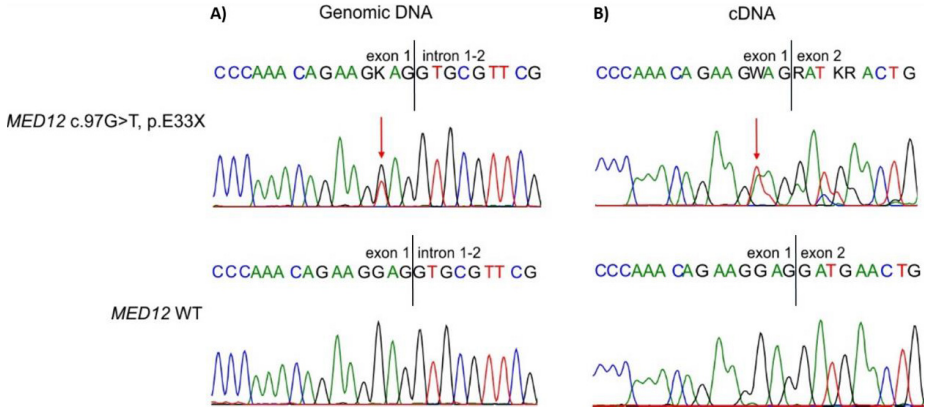

Figure 9. Sequence chromatograms of the *MED12* c.97G>T mutation at the genomic DNA level (A) and at the cDNA level (B). Wild-type reference sequences are shown in the lower lane. The mutant T-allele and the alternative transcript (r.96_99del) are expressed at cDNA level while the wild-type allele is not.

5.1 *MED12* exon 1 nonsense mutation escapes nonsense mediated mRNA decay and encodes an N-terminally truncated protein

Western blot analysis with antibody against the HA-tag in the C-terminus of each construct showed the MED12 derivative to be expressed from all created cell lines. Molecular weight of the E33X product was, however, smaller compared to other MED12 derivatives, and totally undetectable with the N-terminal anti-MED12 antibody (Figure 10). Mass spectrometry-based peptide identification validated the observation, as no peptides originating from the N-terminus of the protein were identified in the E33X sample. The first MED12 peptide sequence observed in the E33X sample located to amino acids 163-174, and methionine (M) 154 was experimentally verified as the alternative start site for translation (Figure 11 and Supplementary Figure S3 in the original publication V).

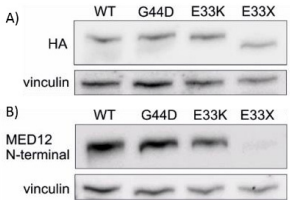

Figure 10. Western blot analyses of total protein isolates from Flp-In 293 T-Rex cells expressing MED12 WT, G44D, E33K, and E33X mutant derivatives. An N-terminally truncated E33X derivative with smaller molecular weight was detected by anti-HA antibody **(A)**. Anti-MED12 antibody, whose epitope resides in the N-terminus of the protein (between aa25-aa75), failed to detect the E33X derivative **(B)**. Vinculin was used as a loading control in both experiments.

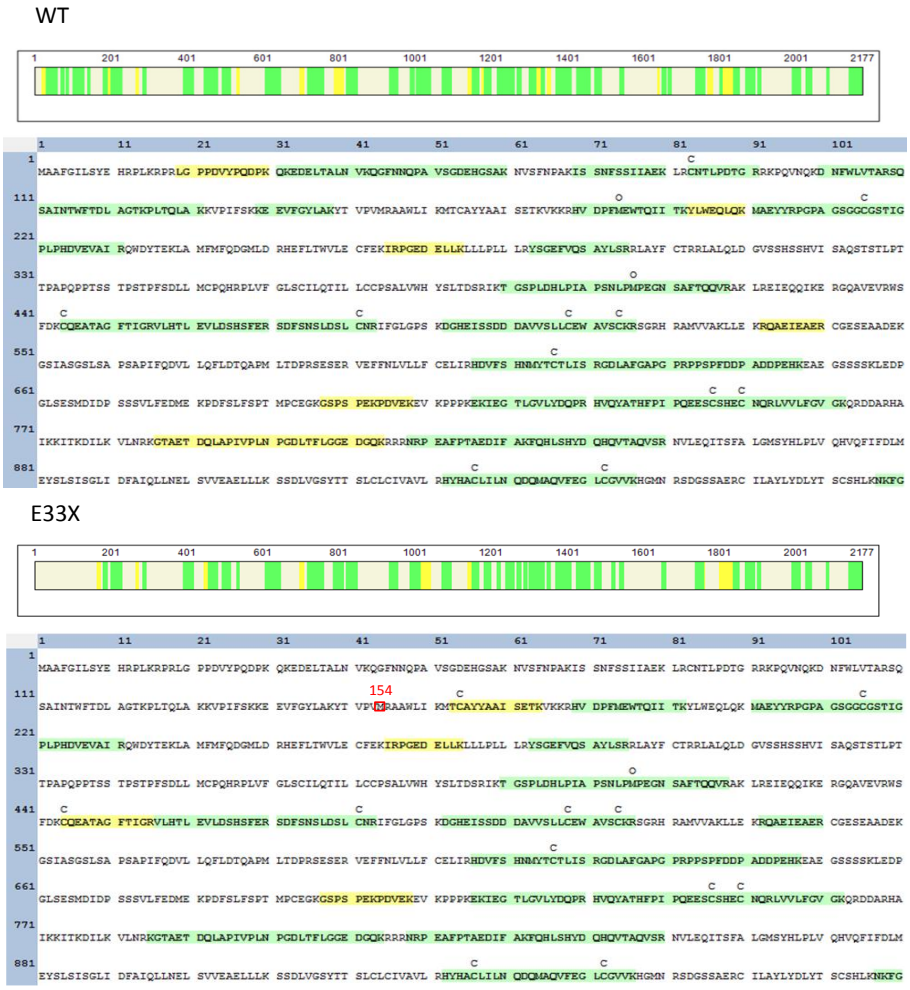

Figure 11. Peptide identification of MED12 WT and E33X derivative utilizing mass spectrometry. The most N-terminal peptide sequence identified in MED12 E33X sample located to amino acids T163-K174. M154 used as alternative start site for translation is marked with red box in the amino acid sequence. Identification of each peptide is indicated as highly confident with a <1% false discovery rate (green) or medium confident (yellow), with a <5% false discovery rate.

5.2 MED12 E33X mutation abolishes the interactions between MED12 and other Mediator components

Affinity purification and subsequent mass spectrometry demonstrated that all interactions between MED12 and other components of the Mediator were lost due to the E33X mutation. The missense mutation E33K, affecting the same amino acid, disrupted only the interactions between MED12 and the kinase module subunits CDK8/19 and Cyclin C. A similar effect with the G44D mutant derivative was observed previously (Turunen *et al.*, 2014) and again in this study. No reduction in the interactions between MED12 and Mediator core components were observed when analyzing E33K and G44D missense mutants (Figure 12).

5.3 MED12 E33X mutant derivative is missing the nuclear localization signal and remains in the cytoplasm

Subcellular localization of MED12 derivatives was analyzed with immunofluorescence. MED12 WT and missense mutants E33K and G44D were localized mainly in the nucleus as expected, whereas the E33X derivative remained in the cytoplasm (Figure 13). Four *in silico* prediction tools predicted an NLS in the N-terminus of the protein between amino acids 13-19, which are strongly evolutionarily conserved. To verify the existence of the NLS, two Flp-In 293 T-Rex cell lines expressing MED12 with modified NLS were created and further analyzed. In constructs NLS1 and NLS2, amino acids 13-16 PLKR and 13-19 PLKRPRL, respectively, were changed to stretches of alanines. Western blot analysis demonstrated the production of full-length MED12 from these constructs. Immunofluorescence analysis showing similar cytoplasmic localization as observed with the E33X derivative verified disabled NLS in both of the constructs (Figure 13). Interactions between the NLS mutant derivatives and all Mediator subunits were severely decreased or almost completely lost in AP-MS analysis (Figure 12).

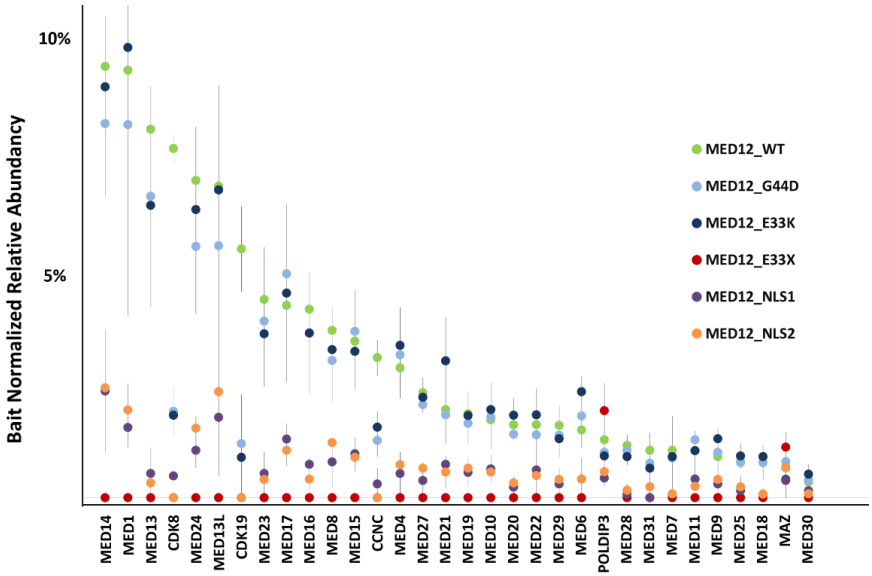

Figure 12. Protein interactions of different MED12 derivatives analyzed by AP-MS. MED12 E33X does not interact with any other component of the Mediator (red dots). Interactions between NLS mutants (purple and orange dots) and Mediator components are decreased, while the missense mutations G44D and E33K (light blue and dark blue dots) abolish only the interactions within the Mediator kinase module when compared to the WT. Error bars represent ± SD.

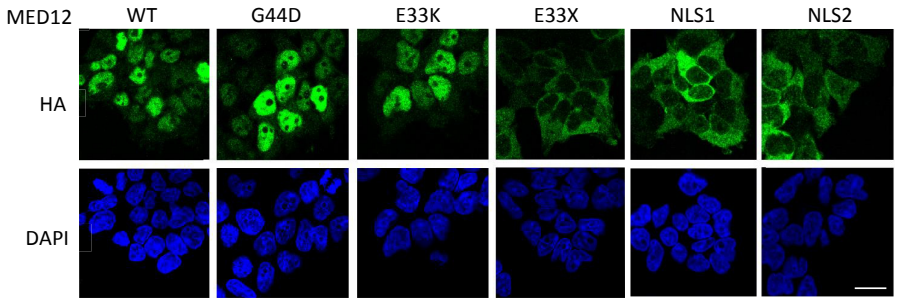

Figure 13. Immunofluorescence analysis of Flp-In 293 T-Rex cells expressing MED12 WT, G44D, E33K, E33X, and NLS1/2 derivatives (original magnification, x 400; scale bar, 20µm). Anti-HA together with fluorescent dye-labeled secondary antibody and DAPI nuclear stain demonstrated the inaccessibility of E33X and NLS1/2 derivatives to the nuclei.

The BioID method was used to analyze the transient protein-protein interactions of MED12 derivatives. Lost or strongly reduced interactions between MED12 E33X/NLS derivatives and all other Mediator components were validated. In addition, analysis revealed that interactions between these derivatives and importin-α proteins (IMA1, IMA4, and IMA7; adaptor proteins binding to an NLS) were severely disrupted. Decreased interactions with the nuclear components of the nuclear pore

complex (NPC; NUP50, NUP153, and TPR) were also observed. Interactions between the corresponding cytoplasmic components of the NPC (RBP2/NUP358, NUP214, and NUP88) and NLS derivatives were increased (Figure 2E in the original publication V). These results suggest that due to the lack of intact NLS, MED12 E33X and NLS derivatives are not recognized and hence localized properly by the nuclear localization machinery.

DISCUSSION

1. The role of *MED12* in uterine leiomyomas

1.1 *MED12* mutations in uterine leiomyomas

MED12, a subunit of the transcription regulating Mediator complex, was convincingly associated with human tumorigenesis when remarkably frequent mutations in the gene were identified in uterine leiomyomas (Mäkinen *et al.*, 2011b). Previously, MED12 had been connected to some cancer-related pathways, for instance canonical Wnt-signaling and Sonic hedgehog signaling, but its role has mainly been recognized in various developmental processes. Several different *MED12* germline mutations in the region coding for the LS- and PQL-domains have been established as molecular determinants behind X-linked intellectual disabilities, conditions such as Opitz-Kaveggia syndrome, Lujan-Fryns syndrome, and Ohdo syndrome (MKB-type) also including specific physical anomalies. Recurrent somatic *MED12* mutations have not been observed previously in any tumor type.

The substantial frequency and highly specific distribution of the mutations locating to a narrow area of the gene provide strong evidence about the contribution of *MED12* in the tumorigenesis of uterine leiomyomas. Since the initial observation of *MED12* exon 2 mutations in uterine leiomyomas from Finnish patients, this finding has been repeatedly validated in several populations and ethnicities comprising Caucasian, African, White and Black American, Asian, Hispanic, Russian, Iranian, and Chinese women (Table 9, Figure 14). In all of these studies, a striking distribution of the mutations was observed, with the vast majority of the mutations affecting codon G44, signifying strong positive selection. The overall *MED12* mutation frequencies have been considerably high in all of these studies, ranging from 50 to 80% in most of them. The discovery of recurrent *MED12* mutations in uterine leiomyomas resulted in a huge leap forward in the molecular research of these extremely common, yet until then very poorly characterized, tumors.

Table 9. *MED12* exon 2 mutations observed in conventional uterine leiomyomas of women from different populations.

Population	Samples analyzed	*MED12* exon 2 mutation-positive samples	Mutation frequency	Reference
Finnish	225	159	71%	(Mäkinen *et al.*, 2011b)
South African	28	14	50%	(Mäkinen *et al.*, 2011a)
German	80	47	59%	(Markowski *et al.*, 2012)
North American	148	100	68%	(McGuire *et al.*, 2012)
Korean	67	35	52%	(Je *et al.*, 2012)
French	9	6	67%	(Perot *et al.*, 2012)
North American	12	9	75%	(Ravegnini *et al.*, 2013)
Finnish	69	41	59%	(Mäkinen *et al.*, 2013)
German	21	10	48%	(Markowski *et al.*, 2013)
Japanese	45	36	80%	(Matsubara *et al.*, 2013)
Dutch	19	11	56%	(de Graaff *et al.*, 2013)
German	12	11	92%	(Rieker *et al.*, 2013)
German	256	179	70%	(Markowski *et al.*, 2014)
North American	28	15	54%	(Schwetye *et al.*, 2014)
North American	177	134	76%	(Bertsch *et al.*, 2014)
Finnish	164	137	84%	(Heinonen *et al.*, 2014)
Chinese and North American	40	30	75%	(Zhang *et al.*, 2014)
South American	143	92	64%	(Halder *et al.*, 2015)
Russian	122	63	52%	(Osinovskaya *et al.*, 2015)
Iranian	23	11	48%	(Shahbazi *et al.*, 2015)
Chinese	181	95	53%	(Wang *et al.*, 2015a)
Iranian	103	32	31%	(Sadeghi *et al.*, 2016)
North American/ Austrian/ French	16	10	63%	(Liegl-Atzwanger *et al.*, 2016)

Highly specific mutations in *MED12* represent the only recurrent somatic small-scale changes observed thus far in uterine leiomyomas. Whole-exome or -genome sequencing of *MED12* mutation-negative and -positive tumors have not revealed any other gene to harbor recurrent somatic point mutations relevant in myomagenesis (Mehine *et al.*, 2013; Mäkinen *et al.*, 2014b). Instead, whole-genome sequencing revealed interconnected complex chromosomal rearrangements in part underlying the cytogenetic aberrations seen in uterine leiomyomas (Mehine *et al.*, 2013). These events resembled chromotripsis, the drastic shattering and random reassembling of affected chromosomes, but mostly in a milder form (Forment *et al.*, 2012; Mehine *et al.*, 2013). These complex rearrangements affected only a subset of the tumors, the majority of them being *MED12* mutation-negative. In general, uterine leiomyomas are considered as rather stable tumors. They display relatively few cytogenetic alterations in addition to the initial factor driving tumorigenesis when compared to the highly complex karyotypes of their malignant counterpart, uterine leiomyosarcoma

(Packenham *et al.*, 1997; Sandberg, 2005a; Sandberg, 2005b; Perot *et al.*, 2012; Mehine *et al.*, 2013; Liegl-Atzwanger *et al.*, 2016). Multiple leiomyomas with different *MED12* mutations, as well as tumors driven by different driver alterations, can co-occur in the same uterus. The synchronous occurrence of these tumors has been observed more often in patients with at least one *MED12* mutation-positive tumor (McGuire *et al.*, 2012; Heinonen *et al.*, 2014; Osinovskaya *et al.*, 2015). Additionally, association between positive *MED12* mutation status and smaller tumor size has been shown (Mäkinen *et al.*, 2011b; Markowski *et al.*, 2012; de Graaff *et al.*, 2013; Heinonen *et al.*, 2014). Correlation between *MED12* mutation status and various other clinical variables and lifestyle factors, for instance age at diagnosis or at hysterectomy, age at menarche, parity, infertility, smoking, or alcohol consumption have not been observed (Je *et al.*, 2012; McGuire *et al.*, 2012; de Graaff *et al.*, 2013; Bertsch *et al.*, 2014; Heinonen *et al.*, 2014; Shahbazi *et al.*, 2015).

1.2 Functional impact of *MED12* exon 1 and 2 mutations

Although the etiological evidence of the role of *MED12* in uterine leiomyomas is comprehensive, the underlying molecular mechanisms by which *MED12* mutations elicit leiomyomagenesis are as yet unclear. The functional mechanisms of *MED12* exon 2 mutations in promoting tumorigenesis have been studied in collaboration with the research groups of professors Jussi Taipale and Thomas Boyer. Exon 2 mutations affecting the three most commonly mutated amino acids, G44, L36, and Q43, were included in the study, and the protein-protein interactions between mutant MED12 derivatives and other components of the Mediator, as well as those with WT MED12, were analyzed (Turunen *et al.*, 2014). A comparison revealed that exon 2 mutations disrupt the interaction between MED12 and Cyclin C, leading to decreased association of MED12 with the Cyclin C-CDK8/19 complex and further to diminished Mediator-associated kinase activity. The binding interface required for the interaction with Cyclin C and for the kinase activity of CDK8 was determined to the first 100 amino acids of MED12. In a recent study, an integrated bioinformatic approach utilizing various *in silico* prediction programs and statistical software was used to evaluate the pathogenity of 14 *MED12* missense mutations (Banaganapalli *et al.*, 2016). All mutations, including four exon 2 missense mutations affecting codons Q43 and G44, were evaluated to be deleterious to the MED12 protein structure and stability. Also this computational analysis demonstrated a decreased binding affinity between MED12 mutants and Cyclin C that likely prevents proper assemblage of the kinase module.

The fact that the genomic region which is highly mutated in uterine leiomyomas spans from intron 1 to exon 2 prompted us to investigate whether mutations would exist also in the first exon of the gene. Five *MED12* exon 1 mutations were identified in conventional uterine leiomyomas as a result of the screening (Figure 14). All changes were small deletions located at the end of exon 1 and retaining the reading frame. Mutations in exon 1 slightly increase the overall *MED12* mutation frequency in uterine

leiomyomas and further underline the central role of *MED12* in their pathogenesis. *MED12* exon 1 mutation-positive uterine leiomyomas displayed expression profiles identical with exon 2 mutation-positive lesions, including the clear overexpression of *RAD51B*. This gene has been identified as the most significantly overexpressed gene in *MED12* exon 2 mutation-positive uterine leiomyomas in our previous study (Mehine *et al.*, 2013). Even though the mechanistic effect of RAD51B in leiomyomagenesis is still obscure, its role seems evident, as it is also the most common translocation partner of *HMGA2*-affecting translocations. The functional impact of *MED12* exon 1 and 2 mutations was further unified as immunoprecipitation and kinase activity assays demonstrated a similar disruption of the MED12 Cyclin C-CDK8/19 interaction and reduced Mediator-associated kinase activity also with *MED12* exon 1 mutant constructs. *MED12* exon 1 mutations with functional effects similar to exon 2 mutations substantiate the notion that the N-terminus of MED12 contains a domain which is needed for the compilation of the kinase module and furthermore for the activation of CDK8. The requirement of MED12 for CDK8 kinase activity has been demonstrated previously, and a portion of CDK8 modules is present in a Mediator-free form in the cells (Knuesel *et al.*, 2009b). Whether the RNA Pol II CTD is a true biological substrate for CDK8 phosphorylation remains elusive, and also other substrates, for instance NOTCH1 and STAT transcription factors, have been implicated for CDK8 kinase activity (Knuesel *et al.*, 2009b; Bancerek *et al.*, 2013; Li *et al.*, 2014).

All studies where *MED12* expression at the mRNA level has been assessed through cDNA sequencing have shown predominant expression of the mutant allele (Mäkinen *et al.*, 2011b; Markowski *et al.*, 2012; McGuire *et al.*, 2012; Perot *et al.*, 2012; Shahbazi *et al.*, 2015; Sadeghi *et al.*, 2016). Immunohistochemical analysis revealed MED12 protein expression in all conventional uterine leiomyomas, regardless of the mutation status, although another study observed decreased expression associated with complex *MED12* mutations (insertions and deletions) (Perot *et al.*, 2012; Bertsch *et al.*, 2014). Further evidence about the role of *MED12* in leiomyomagenesis and its oncogenic nature has emerged from animal experiments. A mouse model conditionally expressing the most common uterine leiomyoma-associated *MED12* mutation, c.131G>A (p.G44D), in uterine mesenchymal cells was shown to develop leiomyoma-like lesions in the uterus (Mittal *et al.*, 2015). The mutation was effective in both the *Med12* WT background as well as in cells where the WT *Med12* was conditionally knocked out. Mere conditional knock out of the WT *Med12* in uterine mesenchymal cells did not induce uterine hyperplasia or leiomyomas, indicating *Med12* c.131G>A as a gain-of-function mutation.

1.3 Mutually exclusive drivers of myomagenesis

To date, at least three distinct uterine leiomyoma subgroups with different oncogenic changes driving tumorigenesis have been recognized. Lesions harboring *MED12* mutations form the largest subgroup, accounting for approximately 60-70% of all

uterine leiomyomas. The second most common driver change is chromosomal aberrations leading to *HMGA2* overexpression, with an estimated proportion of ~10%. Biallelic inactivation of *FH*, usually associated with HLRCC syndrome, covers only a minor fraction of the uterine leiomyomas (Mehine *et al.*, 2014; Markowski *et al.*, 2015). All subgroups display unique, highly characteristic expression profiles (Vanharanta *et al.*, 2006; Hodge *et al.*, 2012; Mehine *et al.*, 2013; Mehine *et al.*, 2014; Mehine *et al.*, 2016), and some diverging clinical and histological features are more often associated with a certain subgroup. These include, for instance, a smaller average size of *MED12* mutation-positive tumors and their tendency to be present as multiple lesions, as well as prominent eosinophilic nucleoli surrounded by clear haloes as a typical histological finding in FH-deficient uterine leiomyomas (Launonen *et al.*, 2001; Mäkinen *et al.*, 2011b; Markowski *et al.*, 2012; Sanz-Ortega *et al.*, 2013; Heinonen *et al.*, 2014; Markowski *et al.*, 2014; Joseph *et al.*, 2015). Alterations affecting the *COL4A5-COL4A6* locus on the X chromosome have also been suggested as driver events in the tumorigenesis of sporadic uterine leiomyomas (Mehine *et al.*, 2013). Tumors harboring *COL4A5-COL4A6* aberrations display characteristic expressional profiles, including clear overexpression of the *IRS4* gene adjacent to *COL4A5,* and cluster together in unsupervised hierarchical clustering analysis (Mehine *et al.*, 2013; Mehine *et al.*, 2016).

MED12 mutations and *HMGA2* overexpression have not been observed to co-occur in uterine leiomyomas, although some cytogenetic aberrations, for example those affecting the long arm of chromosome 7 or the *HMGA1* locus at chromosome 6p, have been identified as co-occurring secondary changes (Markowski *et al.*, 2012; Mehine *et al.*, 2013; Bertsch *et al.*, 2014). Observation of somatic *MED12* mutations instead of biallelic *FH* inactivation in individual uterine leiomyomas from HLRCC patients suggests that these driver alterations are also mutually exclusive in myomagenesis (Mäkinen *et al.*, 2013; Heinonen *et al.*, 2014; Mäkinen *et al.*, 2014b). The role of *MED12* mutations in uterine leiomyomas from HLRCC patients and the mutual exclusiveness of *MED12* mutations and biallelic *FH* inactivation were comprehensively investigated in this study. Analysis of an extensive sample set containing 122 uterine leiomyomas from 27 individuals with HLRCC syndrome and 66 conventional sporadic leiomyomas clearly demonstrated that these two driver events are mutually exclusive in uterine leiomyomas. *MED12* mutations were observed in 7% (9/122) of HLRCC patients' uterine leiomyomas, and none of these displayed biallelic *FH* inactivation. In expression profiling, these tumors clustered together with sporadic *MED12* mutation-positive lesions instead of FH-deficient HLRCC syndrome-associated tumors. Germline *FH* mutation did not affect the expression profiles nor alter the clustering of the normal myometrium tissue samples obtained from HLRCC patients. These findings confirm the presumption that given the high frequency of sporadic uterine leiomyomas in general, and furthermore the proportion of *MED12* driven lesions among these, some would occur also in women with HLRCC syndrome. The proportion of HLRCC patients' lesions harboring *MED12* mutation was, however, surprisingly high in the sample set. This was due to a patient with six *MED12* mutation-positive lesions yet carrying a germline mutation

in *FH*. Further analysis of archival tumor specimens from the patient (24 FFPE specimens from previous myomectomies and from other tumors removed at the hysterectomy), revealed only one FH-deficient HLRCC-associated tumor among numerous sporadic *MED12* mutation-positive leiomyomas. This observed distribution of syndrome-associated and sporadic uterine leiomyomas in a single HLRCC patient is extremely unlikely, and implies the presence of an additional genetic factor that either protects from biallelic *FH* inactivation or predisposes the tissue for somatic *MED12* mutations. All sporadic uterine leiomyomas analyzed in the study were FH-proficient, which is in keeping with the previously observed low frequencies of sporadic FH-deficient lesions (Kiuru *et al.*, 2002; Lehtonen *et al.*, 2004; Harrison *et al.*, 2015).

2. *MED12* mutations in other solid tumors

2.1 *MED12* mutations in uterine leiomyosarcomas

In order to investigate the possible role of *MED12* in other tumors, a comprehensive assemblage of tumor types also originating from the mesenchymal tissue or displaying characteristics similar to uterine leiomyomas (uterine leiomyosarcomas and other sarcomas, gastrointestinal stromal tumors, extrauterine leiomyomas, endometrial polyps, lipomas, ovarian cancers, and breast cancers) or with a connection to MED12-related signaling (acute myeloid leukemias, acute lymphoblastic leukemias, myeloproliferative neoplasms, and colorectal cancers) were collected for this study. The genomic region of exon 2 as well as the preceding intron-exon boundary were analyzed for mutations by direct Sanger sequencing. The observation of five *MED12* exon 2 mutations among the analyzed 1158 tumor samples implicates that despite its central contribution to the development of uterine leiomyomas, MED12 does not possess a major role in the tumorigenesis of the tumor types analyzed here. An exception to this conclusion comes in the form of uterine leiomyosarcoma, where three out of the five mutations, two in primary tumors and one in a metastatic lesion, were identified. Uterine leiomyosarcomas are malignant tumors of the myometrium with a high recurrence rate and strong potential for metastatic dissemination (Major *et al.*, 1993; Gadducci *et al.*, 1996; Giuntoli *et al.*, 2003). Symptoms largely resemble those caused by uterine leiomyomas and usually uterine leiomyosarcoma is diagnosed postoperatively in a pathological evaluation (Leibsohn *et al.*, 1990; Gadducci *et al.*, 1996). Uterine leiomyosarcomas are relatively rare (less than 0.4 cases/100,000 women in a year) and, similar to uterine leiomyomas, their incidence is higher among black women (Brooks *et al.*, 2004; Toro *et al.*, 2006; Koivisto-Korander *et al.*, 2012). In our study, 7% of leimyosarcomas harbored *MED12* mutations, all of them being alterations typically seen in uterine leiomyomas. This observation is in line with other studies reporting *MED12* exon 2 mutations in uterine leiomyosarcomas with frequencies varying from 0 to 30% (Figure 14) (Perot *et al.*, 2012; Je *et al.*, 2012; de Graaff *et al.*, 2013; Matsubara *et al.*, 2013; Ravegnini *et al.*, 2013; Markowski *et al.*,

2013; Schwetye *et al.*, 2014; Bertsch *et al.*, 2014; Zhang *et al.*, 2014; Wang *et al.*, 2015a). Recent WES analysis of uterine leiomyosarcomas identified *MED12* mutations at a frequency of 21%, with *MED12* being the third most commonly mutated gene after *TP53* (33%) and *ATRX* (26%) (Mäkinen *et al.*, 2016). All *MED12* mutations affected mutational hotspot amino acids in exon 2, and no mutations in other parts of the gene were observed.

Benign lesions or hyperplasia of the tissue can act as a precursor for malignant tumor development. This concept is well established, for instance in colorectal cancer, where a high proportion of malignant tumors are preceded by benign adenomatous polyps (Fearon and Vogelstein, 1990). The observation of uterine leiomyoma-linked *MED12* exon 2 mutations in uterine leiomyosarcomas could implicate a similar sequential development path from benign leiomyomas to malignant leiomyosarcoma in some cases. This hypothesis has been suggested previously based on a similar X-chromosome inactivation pattern of some leiomyomas and leiomyosarcomas existing in the same uterus and moreover displaying a microscopically visible morphological transition from benign leiomyoma histology to that of malignant leiomyosarcoma (Zhang *et al.*, 2006). Morphologically benign areas in uterine smooth muscle tumors diagnosed as leiomyosarcomas have mostly represented the histology of cellular leiomyomas or leiomyomas with bizarre nuclei. The innocuousness of these areas was also verified by immunohistochemical analysis showing leiomyoma-like expression profiles of p53, proliferation marker Ki-67, progesterone receptor (PR), and estrogen receptor (ER) (Mittal and Joutovsky, 2007; Mittal *et al.*, 2009). In addition, in some uterine leiomyosarcomas with a benign component, identical *MED12* mutations have been identified in both components, suggesting their common origin (Matsubara *et al.*, 2013). Based on comparative genomic hybridization (CGH) analysis, almost all chromosomal aberrations observed in the area resembling leiomyomas were also present in the representative area of the leiomyosarcoma in addition to usually numerous and complex leiomyosarcoma-specific rearrangements (Mittal *et al.*, 2009). In another study, it was noted that uterine leiomyomas harboring near-complete loss of the chromosome 1 short arm are more often of a cellular histological type, and the expressional profile of two such tumors resembled the expressional profiles of uterine leiomyosarcomas (Christacos *et al.*, 2006).

Our previous analysis of histopathological uterine leiomyoma subtypes revealed recurrent *MED12* exon 2 mutations also in these lesions, albeit with significantly lower frequencies. Only among mitotically active leiomyomas, the mutation frequency (38.5%) was close to that of conventional uterine leiomyomas. In leiomyomas with bizarre nuclei and in tumors of the cellular subtype, the frequencies were substantially lower, 17% and 9%, respectively (Mäkinen *et al.*, 2013). Concordant differences in *MED12* mutation frequencies have been observed in other studies (Perot *et al.*, 2012; Matsubara *et al.*, 2013; Zhang *et al.*, 2014; Liegl-Atzwanger *et al.*, 2016), whereas mutations in *TP53* and *PTEN* deletions, frequent in uterine leiomyosarcomas, are also recurrently observed in uterine leiomyomas with bizarre nuclei and in smooth muscle tumors with uncertain malignant potential (STUMP)

(Zhang *et al.*, 2014). Cytogenetic observations and similar mutation frequencies give additional evidence for the hypothetical continuum from a benign uterine leiomyoma precursor to a malignant tumor. According to current understanding, conventional uterine leiomyomas do not possess malignant potential, and only a minor proportion of the histopathological leiomyoma subtypes, supposedly the cellular subtype or leiomyomas with bizarre nuclei, might develop into malignant leiomyosarcoma. Alternatively, instead of being an initiator in the uterine leiomyosarcoma tumorigenesis, *MED12* mutations can act upon already transformed cells by giving an additional growth advantage and thus leading to clonal expansion. Whereas in conventional uterine leiomyomas MED12 expression was verified at the protein level, in uterine leiomyoma variants with bizarre nuclei, STUMPs, and uterine leiomyosarcomas, MED12 expression was reported to be decreased or missing already at the mRNA level and further abolished at the protein level (Perot *et al.*, 2012). The authors hypothesized that *MED12* could act as a tumor suppressor with attenuated function in leimyomagenesis, and that complete loss of function would be involved with the malignant transformation. Decreased MED12 protein levels in uterine leiomyosarcomas were noted in a subsequent study, although another study demonstrated positive MED12 expression at mRNA level for the majority of uterine leiomyosarcomas (Markowski *et al.*, 2013; Bertsch *et al.*, 2014).

MED12 exon 2 mutations were not seen in any other smooth muscle tumors included in our study. In the literature, individual *MED12* exon 2 mutations have been reported in extrauterine leiomyomas, mainly occurring in the abdominal or peritoneal area (in the smooth muscle tissues deriving from the Müllerian duct), as well as in endometrial polyps (Markowski *et al.*, 2012; Ravegnini *et al.*, 2013; Markowski *et al.*, 2013; de Graaff *et al.*, 2013; Rieker *et al.*, 2013; Schwetye *et al.*, 2014). *MED12* exon 2 mutations have also been observed in a small proportion of STUMPs and only in one extrauterine leiomyosarcoma (located at the thigh/femoral bone) (Perot *et al.*, 2012; Schwetye *et al.*, 2014). *MED12* exon 1 mutations were neither observed in smooth muscle tumors (extrauterine leiomyomas, endometrial polyps, uterine leiomyosarcomas) nor in other sarcomas which were included in the exon 1 mutation screening.

2.2 *MED12* mutations in breast tumors

In this study, no *MED12* exon 2 mutations were observed in breast and ovarian cancers which share the strong hormonal dependency with uterine leiomyomas. Concordantly, *MED12* exon 2 mutations have not been reported in the literature in these tumor types, albeit in breast carcinomas, individual mutations are observed throughout the gene (COSMIC and CBioPortal databases). More recent studies have, however, revealed frequent *MED12* exon 2 mutations in fibroadenomas and phyllodes tumors of the breast (Lim *et al.*, 2014; Cani *et al.*, 2015; Yoshida *et al.*, 2015; Nagasawa *et al.*, 2015; Pfarr *et al.*, 2015; Piscuoglio *et al.*, 2015; Ng *et al.*, 2015; Mishima *et al.*, 2015; Tan *et al.*, 2016; Lien *et al.*, 2016; Yoon *et al.*, 2016). Both the frequency and the

distribution of the mutations have been strikingly similar with those seen in the uterine leiomyomas, strongly suggesting common pathogenic mechanisms in these tumors. Of note, *MED12* exon 1 mutations have not been reported in fibroepithelial tumors of the breast, but the majority of studies have analyzed only the second exon of *MED12* for somatic mutations. Only in a few studies (including eight samples of fibroadenomas and 15 samples of phyllodes tumors), next-generation sequencing (WES or targeted sequencing) has been performed on these samples and thus also exon 1 has been analyzed (Lim *et al.*, 2014; Cani *et al.*, 2015). Fibroepithelial tumors of the breast, mostly comprised by fibroadenomas and to a lesser extent by phyllodes tumors, are hormone-dependent tumors consisting of both stromal and epithelial components (Yang *et al.*, 2014). *MED12* mutations were confirmed to occur only in the former compartment, implicating mammary stromal cells as the origin for the lesions (Lim *et al.*, 2014; Yoshida *et al.*, 2015; Mishima *et al.*, 2015). In addition to the similar mutation spectrum, expression analysis of breast fibroadenomas showed strong correlation between the upregulated genes in uterine leiomyomas and breast fibroadenomas (Mäkinen *et al.*, 2011b; Lim *et al.*, 2014). While fibroadenomas are classified as benign, phyllodes tumors can, instead, be categorized into three separate classes based on their histology: benign (accounting ~60% of the tumors), borderline, and malignant with an increasing tendency to recur (Tan *et al.*, 2012; Yang *et al.*, 2014). A sliding continuum of separating characteristics can be challenging to classify, and transformation of fibroadenomas to malignant phyllodes tumors has been proposed in several studies (Noguchi *et al.*, 1995; Kuijper *et al.*, 2002; Hodges *et al.*, 2009; Abe *et al.*, 2011). Initial studies reported high and similar mutation frequencies in phyllodes tumors regardless of their histological grade (Cani *et al.*, 2015; Mishima *et al.*, 2015; Yoshida *et al.*, 2015). In some studies, especially when analyzing larger sample sets, an inverse correlation between *MED12* exon 2 mutation frequency and increasing degree of malignancy has been noted (Piscuoglio *et al.*, 2015; Pfarr *et al.*, 2015; Ng *et al.*, 2015; Yoon *et al.*, 2016). Positive *MED12* mutation status of the tumor was associated with improved disease-free survival in patients with phyllodes tumors (Ng *et al.*, 2015; Yoon *et al.*, 2016). This observation possibly reflects the role of hormonal regulation in these *MED12* mutation-positive tumors, as dependency on estrogen regulation has been observed to decrease with increasing degree of malignancy in phyllodes tumors (Tse *et al.*, 2002). Additional aberrations in known cancer genes, for instance loss-of-function mutations in *TP53* and *RB1*, have been observed in malignant phyllodes tumors and are presumably driving their malignant progression (Cani *et al.*, 2015).

2.3 *MED12* mutations in other hormone-associated tumors

MED12 mutations have also been observed in other hormone-associated tumor types, such as prostate cancer, endometrial cancer, and adrenocortical carcinomas (Barbieri *et al.*, 2012; Uterine Corpus Endometrioid Carcinoma, TCGA, USA, import from ICGC: COSU419; Assie *et al.*, 2014). In all of these phenotypes, the mutation frequency and the location of the mutations differ substantially from those observed

in uterine leiomyomas and breast fibroepithelial tumors. In prostate cancer, *MED12* mutations were initially observed with a frequency of 4.6% including the recurrent L1224F substitution (5/7) in exon 26 at the leucine-serine-rich LS domain (Barbieri *et al.*, 2012). In addition, missense mutations leading to substitutions D727E and P1310Q, both residing in the same LS domain, were identified in this sample set. Some of the subsequent sequencing studies have validated the finding, albeit with lower mutation frequencies, and identified additional missense mutations (M666V, E1144Q, V1216L, V1220E, V1223L, R1357H, N1845T, and R1899Q) in the region (Figure 14) (Prostate adenocarcinoma, TCGA, US and the Canadian Prostate Cancer Genome Network, import from ICGC: PRAD-US and PRAD-CA; Robinson *et al.*, 2015; Kämpjärvi *et al.*, 2016). One mutation affecting G44, the most frequently affected amino acid in uterine leiomyomas and breast fibroepithelial tumors, was also observed in a sequencing study of metastatic, castration-resistant prostate cancers (Robinson *et al.*, 2015). Expectedly, the functional impact of the prostate cancer-associated MED12 mutations (D727E, L1224F, P1310Q, and R1148H analyzed in Kämpjärvi *et al.*, 2016) clearly differs from the effects seen with N-terminal mutations. Proteome-wide analysis of protein-protein interactions utilizing AP-MS revealed diminished connections between a mutant MED12 derivative (L1224F) and MED1, MED13/MED13L, MED14, MED15, MED17, and MED24. The integrity of the kinase module and its association with the core Mediator, as well as Mediator-associated kinase activity, were further confirmed by immunoprecipitation and kinase activity assays. With prostate cancer-associated MED12 mutant derivatives, changes were not observed in the interaction of MED12 with Cyclin C-CDK8/CDK19 nor in the kinase activity. Additional divergence from the N-terminal *MED12* mutations in uterine leiomyomas and breast fibroepithelial tumors is that in prostate tumors, particularly the epithelial cells instead of stromal cells, harbored the *MED12* mutations (Barbieri *et al.*, 2012). *MED12* has been reported to be overexpressed more frequently in metastatic and recurrent castration-resistant prostate cancers than in androgen-sensitive prostate cancer and benign hyperplasia of the prostate (Shaikhibrahim *et al.*, 2014). Based on immunohistochemical analysis, nuclear accumulation of MED12 correlated with markers of activated transforming growth factor beta (TGF-β) signaling and increased cell proliferation in these more advanced prostate cancer tissues.

In endometrioid endometrial carcinoma, the most common type of endometrial cancer, *MED12* mutations have been reported with a frequency of 5%. The mutations were mostly dispersed along the amino acid sequence: E79 in exon 2, and two other amino acids further down in the amino acid sequence being affected twice, and R521 in exon 11 three times (Uterine Corpus Endometrioid Carcinoma, TCGA, USA, import from ICGC: COSU419). In adrenocortical carcinomas, mutations were observed also with a 5% frequency. No amino acids were affected recurrently, but all observed mutations were located at the C-terminus of the protein, at or adjacent to the defined β-catenin binding site (L1560-G2007) (Assie *et al.*, 2014).

2.4 *MED12* mutations in colorectal cancer

Colorectal cancer constitutes the other tumor type where *MED12* exon 2 mutations were identified in this study. Two mutations, one affecting the most commonly mutated amino acid in uterine leiomyomas, G44, were observed among 392 colorectal cancer samples, leading to a mutation frequency of 0.5%. Mutations in the leiomyoma-linked 5' region of *MED12* have been observed in colorectal cancer with similar frequencies in two contemporary studies (Je *et al.*, 2012; Cancer Genome Atlas Network, 2012). Although in our study *MED12* exon 1 mutations were not observed in colorectal cancer samples, few mutations affecting the end of exon 1 have also been identified in colorectal cancer (COSMIC; Cancer Genome Atlas Network, 2012) (Figure 14). Together, these findings suggest that *MED12* might have a role, yet only a minor one, in the pathogenesis of colorectal cancer. Another subunit of the kinase module, CDK8, is involved in a substantial fraction of colorectal cancers through gene amplifications. Enhanced kinase activity increases β-catenin-driven transcription and is involved in β-catenin-mediated transformation (Firestein *et al.*, 2008). Activation of Wnt/β-catenin signaling associating with mutations in CDK8 kinase module components has been observed also in clear cell renal cell carcinomas (Arai *et al.*, 2014). In addition to mutations observed in Cyclin C encoding *CCNC*, *MED12* harbored non-recurrent mutations in its LS-domain in ~5% of the samples. Of note, in uterine leiomyomas, recurrent mutations or genomic rearrangements affecting the other kinase module subunits (MED13, Cyclin C, and CDK8/19) have not been observed (Mehine *et al.*, 2013; Mäkinen *et al.*, 2014a; Mäkinen *et al.*, 2014b).

3. *MED12* in hematological malignancies

3.1 *MED12* mutations in chronic lymphocytic leukemia

Among the hematological malignancies included in Study I (AML, ALL, and myeloproliferative neoplasms polycythemia vera and essential thrombocytosis), no *MED12* exon 2 mutations were observed. Mutation screening of AML and ALL samples and in addition samples of multiple myeloma and non-Hodgkin lymphoma (432 samples in total) in the study by Je and colleagues also did not reveal mutations in these tumor types (Je *et al.*, 2012). A systematic database search utilizing COSMIC revealed, however, five *MED12* mutations in CLL. Four of these mutations altered amino acids that are recurrently affected in uterine leiomyomas and the fifth was located as well in exon 2 of the gene. Screening of more than 700 CLL samples for *MED12* exon 1 and 2 mutations in this study gave further evidence on the role of *MED12* in CLL. Mutated samples were identified with a frequency of 5.2% (37/709) (Figure 14), implicating *MED12* as a one of the recurrently mutated genes in CLL. Similar to mutations detected in female solid tumors, *MED12* exon 1 and 2 mutations in CLL were missense and insertion/deletion mutations resulting in a substitution of a certain amino acid or adding or skipping of varying number of amino acid residues.

On the other hand, a striking difference in the distribution of mutations at exons 1 and 2 was noted when compared to uterine leiomyomas. In CLL, one third of the mutations occured in exon 1, with the last amino acid of the exon, E33, as a mutational hotspot, while in the uterine leiomyomas analyzed in our previous studies, the great majority of mutations (>95%) were identified in exon 2 (Mäkinen *et al.*, 2011a; Mäkinen *et al.*, 2011b; Mäkinen *et al.*, 2013; Heinonen *et al.*, 2014). Whereas in CLL, mutations affecting the E33 residue were missense mutations, in uterine leiomyomas exon 1 mutations were deletions of variable size, three abolishing the last amino acids of the exon from the transcript. Although in this study *MED12* exon 1 mutations were not observed in any other tumor type in addition to uterine leiomyomas, a few mutations affecting E33 have also been identified in colorectal cancer and bladder cancer (COSMIC; Cancer Genome Atlas Network, 2012).

Prior to this finding, uterine leiomyosarcoma was the only malignant tumor type where *MED12* exon 2 mutations were observed recurrently, making CLL the first extrauterine cancer with recurrent 5' end mutations in *MED12*. CLL is the most common leukemia among the adult population in Western societies. It is more common among men, and indeed, CLL was the first malignancy where specific *MED12* exon 1 and 2 mutations were identified in male patients. The disease is characterized by clonal expansions of mature B cells expressing CD5, CD19, and CD23 cell-surface molecules (Fabbri and Dalla-Favera, 2016). The clinical course and outcome of the disease vary from relatively indolent to aggressive and fatal. Two subclasses based on the somatic hypermutation status of the IGHV genes are distinguished in CLL, and among those, unmutated IGHV status (U-CLL) is associated with shorter treatment-free time and overall survival (Fais *et al.*, 1998; Damle *et al.*, 1999; Hamblin *et al.*, 1999). Expression of ZAP-70, enhancing B-cell receptor signaling, as well as unmethylated state of ZAP-70 are additional well-established markers of poor prognosis in CLL (Chen *et al.*, 2002; Rassenti *et al.*, 2004; Kipps, 2007; Rassenti *et al.*, 2008; Claus *et al.*, 2014). Interestingly, *MED12* mutations observed in our screening were significantly associated with these adverse prognostic markers although an effect on treatment-free time or overall survival was not detected.

Cytogenetic alterations affecting the entire chromosome 12 or the chromosomal regions 13q14, 11q22-23, and 17p13, the latter ones usually associated with inactivation of *ATM* and *TP53*, respectively, are typical for CLL. Large-scale sequencing studies have revealed the mutational burden to be lower in CLL compared to most solid tumors (Vogelstein *et al.*, 2013; Fabbri and Dalla-Favera, 2016). Although several recurrently mutated genes have been identified in CLL, only few reach a frequency of 10%. In addition to *ATM* and *TP53*, the *NOTCH1* transcription factor, and *SF3B1*, involved in 3' mRNA splicing, as well as *BIRC3* and *MYD88* leading to activation of NF-κB signaling when mutated, are among the most commonly mutated genes in CLL (Puente *et al.*, 2011; Quesada *et al.*, 2011; Wang *et al.*, 2011a; Landau *et al.*, 2013; Puente *et al.*, 2015; Landau *et al.*, 2015; Fabbri and Dalla-Favera, 2016; Rossi and Gaidano, 2016). Individual *MED12* mutations in CLL were observed in previous sequencing studies, but with lower frequencies than in our

direct mutation screening (Quesada *et al.*, 2011; Wang *et al.*, 2011a; Schuh *et al.*, 2012; Landau *et al.*, 2013; Damm *et al.*, 2014). The results of our study highlight the importance and utility of somatic mutation databases in identifying novel cancer genes or connecting known cancer genes to new phenotypes when genes are mutated only at a moderate frequency. Currently, *MED12* is recognized as a relevant cancer gene in CLL and has been included in the panel for targeted sequencing of CLL genes (Guieze *et al.*, 2015). In this study, sequencing of relapsed or refractory CLL identified *MED12* mutations with a frequency of 8.8%. This higher mutation frequency was in line with the predominant unmutated IGHV status of the samples. *MED12* exon 1 and 2 mutations have been observed also in more recent WES analyses of CLL samples (Landau *et al.*, 2015; Ljungstrom *et al.*, 2016).

3.1.1 Wnt signaling in uterine leiomyomas and CLL

The observation of very specific *MED12* exon 1 and 2 mutations in clearly distinct tumor types, solid tumors of the uterus and breast, and hematological malignancy of the elderly, is intriguing and calls for further investigation. Dysregulation of the canonical Wnt/β-catenin signaling pathway is seen in a variety of tumors and is one of the pathways implicated in the pathogenesis of *MED12* mutation-associated tumor types. Constitutive activation of β-catenin in the uterine mesenchyme of mice has been shown to induce the formation of mesenchymal tumors which resemble human uterine leiomyomas (Tanwar *et al.*, 2009). Further, Wnt/β-catenin activated transcription was observed to promote the proliferation of human leiomyoma side-population cells in co-culture with mature myometrial cells. Here, the signaling pathway was activated in a paracrine manner by the estrogen/progesterone dependent expression of Wnt-signaling activators in accompanying myometrial cells (Ono *et al.*, 2013). Consistently, inhibition of the Wnt/β-catenin pathway with three distinct small molecule inhibitors resulted in decreased proliferation in primary cultures of leiomyoma cells (Ono *et al.*, 2014b). Altered Wnt signaling has been observed when analyzing the expression profiles of uterine leiomyomas, although in our recent study overexpression of pathway inhibitors was noted (Mäkinen *et al.*, 2011b; Mehine *et al.*, 2016). Upregulated expression of *Wnt4*, a pathway activator that is involved in female sex development, was observed in *MED12* mutation-positive uterine leiomyomas compared to mutation-negative leiomyomas and the myometrium (Markowski *et al.*, 2012). Nuclear accumulation of β-catenin as a result of the activated Wnt pathway was observed in 6/11 *MED12* mutation-positive uterine leiomyomas, while it was absent in all *MED12* mutation-negative tumors (de Graaff *et al.*, 2013). In another study, however, no nuclear accumulation was observed in an analysis of 33 uterine smooth muscle tumors (leiomyomas, leiomyosarcomas and STUMPs), irrespective of their *MED12* mutation status (Perot *et al.*, 2012).

Activation of the Wnt/β-catenin signaling pathway has also been implicated in CLL leukemogenesis. Several factors of the signaling pathway, including Wnt growth factors, frizzled receptors, and LRP co-receptors, as well as the Wnt-regulated

transcription factor LEF1, are overexpressed in CLL cells compared to normal B cells (Lu *et al.*, 2004; Gutierrez *et al.*, 2010; Wang *et al.*, 2014a). Downregulation of the signaling pathway using either small molecule inhibitors or small interfering RNAs (siRNA) decreased the survival of CLL cells (Lu *et al.*, 2004; Gutierrez *et al.*, 2010). In addition, somatic mutations in Wnt pathway members are observed more often than expected in CLL, and some have been shown to promote cell survival through activation of the pathway (Wang *et al.*, 2011b; Wang *et al.*, 2014a).

MED12 constitutes the binding interface for β-catenin in Mediator and is required for correct activation of the Wnt signaling pathway (Kim *et al.*, 2006; Rocha *et al.*, 2010). MED12 has also been implicated in noncanonical Wnt/PCP signaling and was shown to regulate the closure of the neural tube in developing mouse embryos (Rocha *et al.*, 2010). This pathway is also established in the pathogenesis of CLL by regulating the migration and invasion of B cells (Kaucka *et al.*, 2013). Whether specific *MED12* mutations per se affect Wnt signaling in uterine leiomyomas and CLL and the detailed role of this aberrantly regulated pathway in these diseases remain obscure and warrant further investigation.

3.2 *MED12* nonsense mutation in T-cell acute lymphoblastic leukemia

Somatic *MED12* exon 1 and 2 mutations observed recurrently in uterine tumors, breast fibroepithelial tumors, as well as in CLL have not compromised the canonical reading frame of *MED12*. The first nonsense mutation in the region, c.97G>T, leading to a substitution of E33 with a premature stop codon (E33X), was observed in exome sequencing of a T-ALL sample (Kontro *et al.*, 2014). Assays performed with an E33X mutant MED12 construct in our study showed that despite the presence of the nonsense mutation, the protein product was expressed. Further analysis gave evidence that instead of nonsense-mediated mRNA decay (NMD) due to a premature nonsense codon, an alternative translation initiation site is utilized to produce an N-terminally truncated protein. NMD is a translation-dependent process that controls mRNA quality by eliminating erroneous premature termination codon-containing mRNAs. This mechanism is also utilized to maintain cellular homeostasis by downregulating the expression levels of some physiological mRNAs through NMD (Kurosaki and Maquat, 2016). The ability to escape NMD despite nonsense mutations in early exons has been reported in some conditions. For instance, in β-thalassemia, nonsense mutations in β-globin exon 1 (at the 5' end) are able to bypass NMD and to use M55 as an alternative translation initiation site. The resulting N-terminally truncated products do not cause symptoms in heterozygous carriers, whereas NMD-escaping nonsense mutations at the 3' end of the gene result in C-terminally truncated products that lead to a rare symptomatic heterozygous form of β-thalassemia (Neu-Yilik *et al.*, 2011). In the case of the *MED12* exon 1 nonsense mutation, a similar mechanism is observed in a somatic context. In a recent study, this phenomenon was analyzed systematically in the TCGA tumor data and was observed to be a relatively common feature in cancer (Lindeboom *et al.*, 2016).

In addition to a differing N-terminal amino acid composition, the subcellular localization of the truncated protein was altered from the nucleus to the cytoplasm. A nuclear localization signal (NLS) was predicted to be localized in the lost region and was experimentally verified using NLS-modified constructs (Figure 14). Protein-protein interactions between MED12 derivatives with a missing or manipulated NLS and other components of the Mediator, in both the kinase module and the core, were severely impaired. A coherent decrease in the interactions is most likely due to the mislocalization of the mutant proteins, although an effect as a result of a missing or altered binding interface at the MED12 N-terminus cannot be excluded. A nuclear transportation mechanism where α importins recognize the NLS and direct the protein to the cytoplasmic side of the nuclear pore complex (NPC), leading to NPC central channel entry and subsequent localization to the nucleus, is suggested by the observed interaction changes between MED12 derivatives and NPC components. In addition to the established nuclear role of MED12 as a part of the Mediator, MED12 has been shown to function in the cytoplasm. Via a physical interaction, cytoplasmic MED12 negatively regulates the proper maturation of transforming growth factor beta receptor 2 (TGF-βR2) and its expression on the cell surface (Huang *et al.*, 2012). Loss of MED12 expression and subsequent activation of canonical TGF-β as well as MEK/ERK signaling has been shown to confer resistance against various targeted cancer drugs and chemotherapy in a number of cancer cells (Huang *et al.*, 2012; Brunen *et al.*, 2013). Aberrant activation of TGF-β signaling protects cells from chemotherapy-induced apoptosis, and both pathways induce the epithelial-to-mesenchymal transition, promoting tumorigenesis. Interaction between MED12 and TGF-βR2 was not seen in our MS data, which might, however, be due to the modified construct or the type of cell line used in the experiments. MED12 methylation has also been implicated in chemotherapy response. CARM1 arginine methyltransferase was shown to methylate two C-terminal arginines of MED12 and thus sensitizes breast cancer cells to commonly used chemotherapy drugs (Wang *et al.*, 2015b).

Although the impact of *MED12* nonsense mutation on the development of T-ALL cannot be concluded based on a single case with several other mutations in leukemia associated genes (*NOTCH1*, *KRAS*, *SUZ12*, *KDM6A*, and *STAT5B*), the production of an N-terminally truncated protein as a result of this highly unusual mutation suggests an important role for MED12 in normal cell functions. The mutation was present in the diagnostic sample as well as in the sample taken at the time of relapse, further emphasizing its relevance in disease progression. Of note, the kinase module components Cyclin C and CDK8/19 have been previously associated with T-ALL tumorigenesis. The Cyclin C-CDK8/19 complex suppresses oncogenic NOTCH1 signaling by phosphorylation and subsequent degradation of its intracellular domain (ICN1) (Li *et al.*, 2014). Deletions affecting Cyclin C, as well as mutations in the ICN1 domain, disrupt this tumor-suppressive mechanism and are frequently seen in T-ALL.

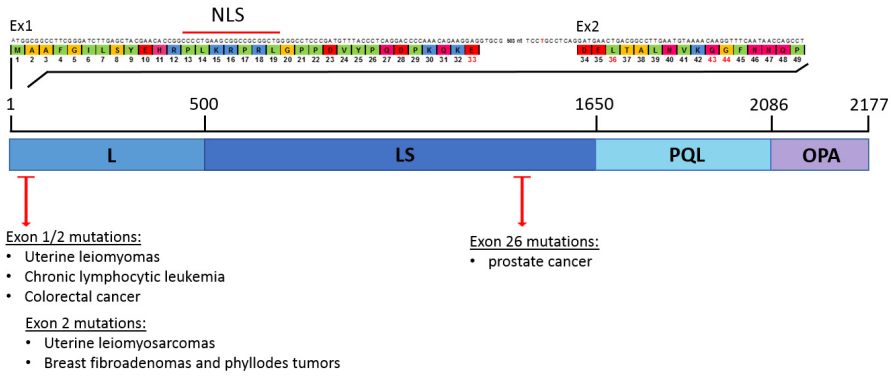

Figure 14. Schematic diagram of MED12 with recurrently observed somatic mutations. Different tumor types where somatic *MED12* mutations are observed recurrently are indicated below the protein domain alignment. Most somatic mutations are located at a highly conserved region consisting the end of exon 1, the intron-exon boundary in the first intron, as well as the beginning of exon 2. Codons and one intronic nucleotide representing the mutational hotspots are marked with red in the amino acid alignment. In prostate cancer individual mutations are also observed between amino acids M666 and R1899. An NLS was identified at the N-terminus of MED12 (marked above the amino acid alignment). Amino acids are colored according to their side-chain pK_as and charge at physiological pH 7.4. Green=hydrophobic; yellow=small non-polar; red=negatively charged; magenta=polar; blue=positively charged. Protein domains: L=leucine-rich; LS=leucine-serine-rich; PQL=proline-glutamine-leucine-rich; OPA=glutamine-rich opposite paired domain.

The relevance of *MED12* mutations in hematological malignancies is emerging as more and more samples representing various phenotypes are being sequenced. Targeted sequencing of T- and NK-cell post-transplant lymphoproliferative diseases revealed *MED12* missense mutations in three of the 18 tumors analyzed. One of the mutations affected L36 in exon 2, and two others in exons 11 and 18 were also predicted as potentially damaging (Margolskee *et al.*, 2016). *MED12* mutations have recently been observed also in pediatric T-ALL (Spinella *et al.*, 2016; Liu *et al.*, 2015, American Society of Hematology 57th Annual Meeting 2015, abstract #691). In addition, MED12 was very recently identified as an essential regulator of the hematopoietic stem cell homeostasis (Aranda-Orgilles *et al.*, 2016). The protein was observed to localize mainly to the hematopoietic stem and progenitor cell-specific enhancer regions, and furthermore to colocalize with essential hematopoietic transcription factors such as RUNX1 and GATA2. *Med12* deletion caused a decrease in hematopoietic stem cell expression signatures and led to a rapid bone marrow failure and lethality. The effect was independent from the association with the kinase module as MED13, Cyclin C, and CDK8 were dispensable for the maintenance of hematopoiesis. As MED12 function is essential for physiological hematopoiesis, it would seem reasonable to consider putative gain-of-function *MED12* mutations observed in hematological malignancies as drivers of leukemogenesis.

CONCLUSIONS AND FUTURE PROSPECTS

MED12, a subunit of the Mediator complex and an important regulator of transcription, was recently identified as a key driver of leiomyomagenesis, as highly specific missense mutations and in-frame insertions and deletions affecting exon 2 of the gene were observed in 70% of uterine leiomyomas. Later it was shown that these mutations disrupt the integrity of the Mediator kinase module and abolish its kinase activity. This thesis work was conducted to further characterize the role of *MED12* in the tumorigenesis of uterine leiomyomas and to analyze its contribution to the development of various other human tumors. Frame-retaining deletions affecting the first exon of *MED12* in uterine leiomyomas were identified in this study. Exon 1 mutations had similar effects on the protein-protein interactions and the tumors' gene expression profiles as was previously observed in the context of exon 2 mutations. These results specify the domain in MED12 which is needed for the compilation of the Mediator kinase module and further highlight the role of *MED12* in the tumorigenesis of these extremely common female tumors. *MED12* mutations have been demonstrated as mutually exclusive with chromosomal rearrangements leading to the overexpression of *HMGA2*, a driver event observed in ~10% of uterine leiomyomas. Here, analysis of hereditary uterine leiomyomatosis and renal cell cancer syndrome patients' uterine leiomyomas clearly indicated that *MED12* mutations and biallelic *FH* inactivation are also mutually exclusive as drivers of leiomyomagenesis. In addition to syndrome-associated FH-deficient tumors, *MED12*-driven sporadic uterine leiomyomas do infrequently occur in HLRCC patients' uteri, whereas outside the context of HLRCC, tumors driven by two somatic mutations in *FH* are very rare.

As a result of the extensive mutation screenings performed in this study, recurrent *MED12* exon 2 mutations were identified in uterine leiomyosarcoma. This observation and the recently reported identification of frequent *MED12* exon 2 mutations in breast fibroepithelial tumors emphasize the gene's role in hormone-dependent female tumors. This finding also demonstrates that *MED12* mutations are not restricted to benign tumors and gives further strength to the previously suggested possibility that some uterine leiomyomas act as precursors for malignant leiomyosarcomas. Additional work is needed to elucidate the putative association of *MED12* mutations and hormonal regulation and to further investigate the molecular mechanisms in the progression of uterine leiomyosarcomas. As aggressive uterine leiomyosarcomas are usually diagnosed postoperatively, identification of benign uterine leiomyoma precursors with malignant potential, if they exist, would be highly important. In addition to identifying suitable molecular markers for their detection, a feasible sampling method is also required. Analysis of cell-free DNA from the circulating blood has been promising in clinical use and could be applicable also in this context.

Intriguingly, CLL, the most common hematological disease of adults, was the other malignancy where recurrent *MED12* mutations were observed in this study. Mutations

consisted of missense mutations and small insertions and deletions similar to those observed in solid female tumors, but their distribution to the site of exon 1 and exon 2 differed significantly between the tumor types. Whether this discrepancy represents a true tissue-specific functional difference needs to be studied more thoroughly with cell type-specific experiments. Also, the observed associations between positive *MED12* mutation status and several markers of poor prognosis requires validation. Analysis of a larger sample material from high-risk patients could reveal differences also in clinical outcomes that, despite strong associations, were not observed in our study.

Functional assessment of the first *MED12* exon 1 nonsense mutation observed in T-ALL identified an NLS at the beginning of MED12 and demonstrated it to be required for the proper nuclear localization of the protein. The protein was expressed despite the mutation using an alternative translation initiation site, leading to an N-terminally truncated product that remained in the cytoplasm. These results suggest cytoplasmic interactions for MED12 that are important for cell viability and highlight the utility of even individual mutations in studying the structure and normal functions of proteins.

Frequent oncogenic-like mutations in *MED12* pose an attractive target for the treatment of *MED12* mutation-associated tumors in the future. More detailed knowledge about the exact molecular mechanisms caused by these specific mutations in different tissues is needed to understand the process of *MED12*-driven tumorigenesis and to identify optimal targets for treatment in a given tumor type. Success of such targeted treatment in the case of uterine leiomyomas might, however, be hampered by the synchronous occurrence of multiple tumors, possibly driven by different genetic aberrations, in the uterus. Advances in the molecular classification of uterine leiomyomas and accumulating information about their clinical features should enable precise characterization of individual lesions and provide tools for more accurate diagnostics and development of personalized treatments.

ACKNOWLEDGEMENTS

This study was carried out at the Department of Medical and Clinical Genetics, Medicum and Genome-Scale Biology Research Program, Research Programs Unit, Faculty of Medicine, University of Helsinki during 2012-2016. Present and former directors are warmly thanked for providing excellent research facilities and an inspiring working environment.

For financial support, I gratefully acknowledge the Helsinki Graduate Program in Biotechnology and Molecular Biology (GPBM) and the Integrative Life Science (ILS) Doctoral Program, the Biomedicum Helsinki Foundation, the Cancer Society of Finland, the Emil Aaltonen Foundation, the K. Albin Johansson Foundation, the Maud Kuistila Memorial Foundation, and the Orion-Farmos Research Foundation.

I would like to express my sincere gratitude to my supervisor Pia Vahteristo, who has been a true role model for me scientifically and also in other areas of life. I highly value your expertise and your enthusiasm towards science, and I hope that I have managed to absorb some of your kindness, your never-ending positivity, and your commitment to both work and family during these years. Thank you for your guidance and for always believing in me. I am also most grateful to my other supervisor Lauri Aaltonen for the opportunity to work in his research group, which is beyond compare. Thank you for the invaluable experience that I have gained working under the supervision of such a great and devoted scientist that you are. Thank you for showing that no goal is too high, as long as there is enough coffee to get there.

I wish to thank warmly my thesis committee members Miina Ollikainen and Oskari Heikinheimo for their encouragement and valuable support during my doctoral studies. Also, Miina Ollikainen and Katri Pylkäs are sincerely thanked for reviewing my thesis; your comments and the discussions we had are greatly appreciated. I would also like to extend my gratitude to Diana Cousminer for the language review of the thesis.

I wish to sincerely acknowledge all collaborators who have provided samples, data, and their technical and scientific expertise, and without whom this work would not have been possible: Johanna Arola, Jan Böhm, Omar Abdel-Wahab, Liisa Pelttari, Heinrich Schrewe, Heli Nevanlinna, Ross Levine, Peter Hokland, Tom Böhling, Jukka-Pekka Mecklin, Ralf Bützow, Min Ju Park, Nam Hee Kim, Alison Clark, Heikki Järvinen, Ian Tomlinson, Zephne van der Spuy, Jari Sjöberg, Thomas Boyer, Salla Välipakka, Outi Uimari, Anne Ahtikoski, Norma Frizzell, Nanna Sarvilinna, Tiina Järvinen, Amy Ruppert, Leigha Senter, Kevin Hoag, Olli Dufva, Mika Kontro, Laura Rassenti, Erin Hertlein, Thomas Kipps, Kimmo Porkka, John Byrd, Albert de la Chapelle, Salla Keskitalo, Pernilla von Nandelstadh, Xiaonan Liu, Ville Rantanen, Matias Kinnunen, Heikki Kuusanmäki, Mikko Turunen, Jussi Taipale, Caroline Heckman, Kaisa Lehti, Satu Mustjoki, and Markku Varjosalo. The Genomics Unit of Technology Centre, Institute for Molecular Medicine Finland, the Biomedicum Functional Genomics Unit, and the Biomedicum Imaging Unit are kindly thanked for their excellent services.

It has been a great pleasure to work with all the members of the Aaltonen group. Present and former post-docs Auli, Outi, Sari, Heli, Rainer, Javier, Esa, Mervi, Eevi, Alexandra, Niko, Miia, Kimmo, Linda vdB., Netta, Linda F., and Simona are warmly thanked for all the help and guidance they have provided me with during my PhD project. Auli is dearly thanked for sharing her vast experience especially in the field of cell work with me, and Esa for his friendship and patience explaining the secrets of statistics to me. Sari is thanked for being such a great roommate and traveling companion with whom I could share so much regarding work and also the childish life. Mervi and Eevi are thanked for all the scientific and otherwise deep conversations that we have had, and for all the fun games and presentations you have organized for the lab! Netta, thank you for guiding me into the world of uterine leiomyomas and for all the instructions and advice you have given me during these years about the sample collections, lab protocols, data analysis, applying funding… you name it! Thank you for your company in the meetings

and excellent advance planning regarding the trips, for our seamless collaboration in many projects, and our friendship in and outside the lab. Outi, thank you for our friendship that started at the HUGO meeting in 2006 and has lasted ever since despite the different countries and continents where we have lived during these years. I admire your scientific expertise and vision and am grateful for all the help and guidance I have gotten from you. We share so much (starting from our nursing history), that I predict, and hope for, many more decades of friendship for us.

I have been privileged to work with excellent technical personnel: Sini N., Inga-Lill, Iina V., Sirpa, Marjo, Mairi, Alison, Sini K., Heikki M., Jiri, and Lauri S. are all thanked for their technical expertise and willingness to help with all the issues regarding the samples, patient information, experiments, orders and other practicalities, programs, data management, etc. Sini N. and Iina are especially thanked for organizing so many fun events for the group. My heartfelt thanks go to fellow PhD students in the Aaltonen lab during these years: Silva, Johanna, Miika, Riku, Iikki, Hanna, Tatiana, Ulrika, Heikki R., Tomas, Hande, and Jaana, and also to those who visited the Aaltonen group during these years. Special thanks go to the members of the myoma team: Miika, Hanna, and Jaana, as well as post-docs Netta, Eevi, and Simona. Thank you for the brilliant team work and for all the inspiring conversations we have had. I wish to thank Miika, "Mr. Myoma", for his central contribution to myoma studies and for all the help he has given me with the data analysis, and Netta and Hanna for their excellent management of the myoma samples and all the associated data. Johanna is thanked for her trustworthiness and friendship also outside the lab. Thank you all for the enjoyable conversations during lunch breaks and hilarious moments at the coffee corner and parties; laughing makes you live longer!

I would like to express my heartfelt thanks to all the members of the KiVa lab: Tuomas, Annukka, Terhi, Anna, Elina, Saara, and Janne. Annukka is especially thanked for lightning up the early mornings in the office and Terhi and Elina for the fun trip to EACR meeting in Manchester. Tuomas, thank you for the close collaboration during the recent years and for all the support and encouragement you have given me.

My sincere thanks go to my wonderful friends outside work. The Vääksy girls and boys: Terhi, Mari, Hanna, Sanna, Minna A., Minna G., Monna, Mikko Y-P., and Mikko N. are thanked for millions of unforgettable moments we have shared around the Europe. I value dearly our lifelong friendship and your endless support. I wish to thank all my friends from the University of Turku: Ville H., Pipsa, Kaisa, Netta, Susanna, Ville V., Minja, and Johanna, as well as Nevanlinnan neidot: Outi, Anitta, Kirsi, and Reetta for all the fun moments when studying/working together and for your cheerful company at our get-togethers.

Finally, I would like to thank my family; mum, dad, my sisters Maiju and Tuulia with their families, and also my family-in-law for all their love and support. I am grateful to my parents Hanna and Seppo for their never-ending love and care and for always believing in me. Maiju and Tuulia are thanked for all their opinions (asked and not asked for) as well as all their appreciative support. Thank you all for the memorable days filled with laughter and joy that we have spent together.

Above all, I want to thank my husband Petteri and our dear children. Sulo and Roosa, thank you for all the joy and spark you have brought into my life and for the unconditional love you pour on me, especially when I deserve it the least. Petteri, thank you for everything. You make my world go around. I love you all to the moon and back.

I express my deepest gratitude to the patients and families who have participated in this study.

Helsinki, October 2016

Kati Kämpjärvi

REFERENCES

Aaltonen LA, Salovaara R, Kristo P, Canzian F, Hemminki A, Peltomäki P, Chadwick RB, Kaariainen H, Eskelinen M, Järvinen H, Mecklin JP, de la Chapelle A (1998) Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med* **338**: 1481-1487.

Abe M, Miyata S, Nishimura S, Iijima K, Makita M, Akiyama F, Iwase T (2011) Malignant transformation of breast fibroadenoma to malignant phyllodes tumor: long-term outcome of 36 malignant phyllodes tumors. *Breast Cancer* **18**: 268-272.

Adam J, Hatipoglu E, O'Flaherty L, Ternette N, Sahgal N, Lockstone H, Baban D, Nye E, Stamp GW, Wolhuter K, Stevens M, Fischer R, Carmeliet P, Maxwell PH, Pugh CW, Frizzell N, Soga T, Kessler BM, El-Bahrawy M, Ratcliffe PJ, Pollard PJ (2011) Renal cyst formation in Fh1-deficient mice is independent of the Hif/Phd pathway: roles for fumarate in KEAP1 succination and Nrf2 signaling. *Cancer Cell* **20**: 524-537.

Advani AS and Pendergast AM (2002) Bcr-Abl variants: biological and clinical aspects. *Leuk Res* **26**: 713-720.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. *Nat Methods* **7**: 248-249.

Ahvenainen T, Lehtonen HJ, Lehtonen R, Vahteristo P, Aittomäki K, Baynam G, Dommering C, Eng C, Gruber SB, Gronberg H, Harvima R, Herva R, Hietala M, Kujala M, Kaariainen H, Sunde L, Vierimaa O, Pollard PJ, Tomlinson IP, Bjorck E, Aaltonen LA, Launonen V (2008) Mutation screening of fumarate hydratase by multiplex ligation-dependent probe amplification: detection of exonic deletion in a patient with leiomyomatosis and renal cell cancer. *Cancer Genet Cytogenet* **183**: 83-88.

Akoulitchev S, Chuikov S, Reinberg D (2000) TFIIH is negatively regulated by cdk8-containing mediator complexes. *Nature* **407**: 102-106.

Alam NA, Barclay E, Rowan AJ, Tyrer JP, Calonje E, Manek S, Kelsell D, Leigh I, Olpin S, Tomlinson IP (2005a) Clinical features of multiple cutaneous and uterine leiomyomatosis: an underdiagnosed tumor syndrome. *Arch Dermatol* **141**: 199-206.

Alam NA, Bevan S, Churchman M, Barclay E, Barker K, Jaeger EE, Nelson HM, Healy E, Pembroke AC, Friedmann PS, Dalziel K, Calonje E, Anderson J, August PJ, Davies MG, Felix R, Munro CS, Murdoch M, Rendall J, Kennedy S, Leigh IM, Kelsell DP, Tomlinson IP, Houlston RS (2001) Localization of a gene (MCUL1) for multiple cutaneous leiomyomata and uterine fibroids to chromosome 1q42.3-q43. *Am J Hum Genet* **68**: 1264-1269.

Alam NA, Olpin S, Leigh IM (2005b) Fumarate hydratase mutations and predisposition to cutaneous leiomyomas, uterine leiomyomas and renal cancer. *Br J Dermatol* **153**: 11-17.

Alam NA, Rowan AJ, Wortham NC, Pollard PJ, Mitchell M, Tyrer JP, Barclay E, Calonje E, Manek S, Adams SJ, Bowers PW, Burrows NP, Charles-Holmes R, Cook LJ, Daly BM, Ford GP, Fuller LC, Hadfield-Jones SE, Hardwick N, Highet AS, Keefe M, MacDonald-Hull SP, Potts ED, Crone M, Wilkinson S, Camacho-Martinez F, Jablonska S, Ratnavel R, MacDonald A, Mann RJ, Grice K, Guillet G, Lewis-Jones MS, McGrath H, Seukeran DC, Morrison PJ, Fleming S, Rahman S, Kelsell D, Leigh I, Olpin S, Tomlinson IP (2003) Genetic and functional analyses of FH mutations in multiple cutaneous and uterine leiomyomatosis, hereditary leiomyomatosis and renal cancer, and fumarate hydratase deficiency. *Hum Mol Genet* **12**: 1241-1252.

Albala JS, Thelen MP, Prange C, Fan W, Christensen M, Thompson LH, Lennon GG (1997) Identification of a novel human RAD51 homolog, RAD51B. *Genomics* **46**: 476-479.

Alimonti A, Carracedo A, Clohessy JG, Trotman LC, Nardella C, Egia A, Salmena L, Sampieri K, Haveman WJ, Brogi E, Richardson AL, Zhang J, Pandolfi PP (2010) Subtle variations in Pten dose determine cancer susceptibility. *Nat Genet* **42**: 454-458.

Allen BL and Taatjes DJ (2015) The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol* **16**: 155-166.

Alsolami S, El-Bahrawy M, Kalloger SE, AlDaoud N, Pathak TB, Chung CT, Mulligan AM, Tomlinson IP, Pollard PJ, Gilks CB, McCluggage WG, Clarke BA (2014) Current morphologic criteria perform poorly in identifying hereditary leiomyomatosis and renal cell carcinoma syndrome-associated uterine leiomyomas. *Int J Gynecol Pathol* **33**: 560-567.

Antoniou AC, Spurdle AB, Sinilnikova OM, Healey S, Pooley KA, Schmutzler RK, Versmold B, Engel C, Meindl A, Arnold N, Hofmann W, Sutter C, Niederacher D, Deissler H, Caldes T, Kämpjärvi K, Nevanlinna H, Simard J, Beesley J, Chen X, Kathleen Cuningham Consortium for Research into Familial Breast Cancer, Neuhausen SL, Rebbeck TR, Wagner T, Lynch HT, Isaacs C, Weitzel J, Ganz PA, Daly MB, Tomlinson G, Olopade OI, Blum JL, Couch FJ, Peterlongo P, Manoukian S, Barile M, Radice P, Szabo CI, Pereira LH, Greene MH, Rennert G, Lejbkowicz F, Barnett-Griness O, Andrulis IL, Ozcelik H, OCGN, Gerdes AM, Caligo MA, Laitman Y, Kaufman B, Milgrom R, Friedman E, Swedish BRCA1 and BRCA2 study collaborators, Domchek SM, Nathanson KL, Osorio A, Llort G, Milne RL, Benitez J, Hamann U, Hogervorst FB, Manders P, Ligtenberg MJ, van den Ouweland AM, DNA-HEBON collaborators, Peock S, Cook M, Platte R, Evans DG, Eeles R, Pichert G, Chu C, Eccles D, Davidson R, Douglas F, EMBRACE, Godwin AK, Barjhoux L, Mazoyer S, Sobol H, Bourdon V, Eisinger F, Chompret A, Capoulade C, Bressac-de Paillerets B, Lenoir GM, Gauthier-Villars M, Houdayer C, Stoppa-Lyonnet D, GEMO, Chenevix-Trench G, Easton DF, CIMBA (2008) Common breast cancer-predisposition alleles are associated with breast cancer risk in BRCA1 and BRCA2 mutation carriers. *Am J Hum Genet* **82**: 937-948.

Arafah BM and Nasrallah MP (2001) Pituitary tumors: pathophysiology, clinical manifestations and management. *Endocr Relat Cancer* **8**: 287-305.

Arai E, Sakamoto H, Ichikawa H, Totsuka H, Chiku S, Gotoh M, Mori T, Nakatani T, Ohnami S, Nakagawa T, Fujimoto H, Wang L, Aburatani H, Yoshida T, Kanai Y (2014) Multilayer-omics analysis of renal cell carcinoma, including the whole exome, methylome and transcriptome. *Int J Cancer* **135**: 1330-1342.

Aranda-Orgilles B, Saldana-Meyer R, Wang E, Trompouki E, Fassl A, Lau S, Mullenders J, Rocha PP, Raviram R, Guillamot M, Sanchez-Diaz M, Wang K, Kayembe C, Zhang N, Amoasii L, Choudhuri A, Skok JA, Schober M, Reinberg D, Sicinski P, Schrewe H, Tsirigos A, Zon LI, Aifantis I (2016) MED12 Regulates HSC-Specific Enhancers Independently of Mediator Kinase Activity to Control Hematopoiesis. *Cell Stem Cell* Epub ahead of print.

Assie G, Letouze E, Fassnacht M, Jouinot A, Luscap W, Barreau O, Omeiri H, Rodriguez S, Perlemoine K, Rene-Corail F, Elarouci N, Sbiera S, Kroiss M, Allolio B, Waldmann J, Quinkler M, Mannelli M, Mantero F, Papathomas T, De Krijger R, Tabarin A, Kerlan V, Baudin E, Tissier F, Dousset B, Groussin L, Amar L, Clauser E, Bertagna X, Ragazzon B, Beuschlein F, Libe R, de Reynies A, Bertherat J (2014) Integrated genomic characterization of adrenocortical carcinoma. *Nat Genet* **46**: 607-612.

Asturias FJ, Jiang YW, Myers LC, Gustafsson CM, Kornberg RD (1999) Conserved structures of mediator and RNA polymerase II holoenzyme. *Science* **283**: 985-987.

Bajekal N and Li TC (2000) Fibroids, infertility and pregnancy wastage. *Hum Reprod Update* **6**: 614-620.

Banaganapalli B, Mohammed K, Khan IA, Al-Aama JY, Elango R, Shaik NA (2016) A Computational Protein Phenotype Prediction Approach to Analyze the Deleterious Mutations of Human MED12 Gene. *J Cell Biochem* **117**: 2023-2035.

Bancerek J, Poss ZC, Steinparzer I, Sedlyarov V, Pfaffenwimmer T, Mikulic I, Dolken L, Strobl B, Muller M, Taatjes DJ, Kovarik P (2013) CDK8 kinase phosphorylates transcription factor STAT1 to selectively regulate the interferon response. *Immunity* **38**: 250-262.

Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, White TA, Stojanov P, Van Allen E, Stransky N, Nickerson E, Chae SS, Boysen G, Auclair D, Onofrio RC, Park K, Kitabayashi N, MacDonald TY, Sheikh K, Vuong T, Guiducci C, Cibulskis K, Sivachenko A, Carter SL, Saksena G, Voet D, Hussain WM, Ramos AH, Winckler W, Redman MC, Ardlie K, Tewari AK, Mosquera JM, Rupp N, Wild PJ, Moch H, Morrissey C, Nelson PS, Kantoff PW, Gabriel SB, Golub TR, Meyerson M, Lander ES, Getz G, Rubin MA, Garraway LA (2012) Exome

sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* **44**: 685-689.

Bardella C, El-Bahrawy M, Frizzell N, Adam J, Ternette N, Hatipoglu E, Howarth K, O'Flaherty L, Roberts I, Turner G, Taylor J, Giaslakiotis K, Macaulay VM, Harris AL, Chandra A, Lehtonen HJ, Launonen V, Aaltonen LA, Pugh CW, Mihai R, Trudgian D, Kessler B, Baynes JW, Ratcliffe PJ, Tomlinson IP, Pollard PJ (2011) Aberrant succination of proteins in fumarate hydratase-deficient mice and HLRCC patients is a robust biomarker of mutation status. *J Pathol* **225**: 4-11.

Bayley JP, Launonen V, Tomlinson IP (2008) The FH mutation database: an online database of fumarate hydratase mutations involved in the MCUL (HLRCC) tumor syndrome and congenital fumarase deficiency. *BMC Med Genet* **9**: 20-2350-9-20.

Berger AH, Knudson AG, Pandolfi PP (2011) A continuum model for tumour suppression. *Nature* **476**: 163-169.

Bertsch E, Qiang W, Zhang Q, Espona-Fiedler M, Druschitz S, Liu Y, Mittal K, Kong B, Kurita T, Wei JJ (2014) MED12 and HMGA2 mutations: two independent genetic events in uterine leiomyoma and leiomyosarcoma. *Mod Pathol* **27**: 1144-1153.

Boghosian L, Dal Cin P, Sandberg AA (1988) An interstitial deletion of chromosome 7 may characterize a subgroup of uterine leiomyoma. *Cancer Genet Cytogenet* **34**: 207-208.

Boland CR and Goel A (2010) Microsatellite instability in colorectal cancer. *Gastroenterology* **138**: 2073-2087.e3.

Borggrefe T, Davis R, Erdjument-Bromage H, Tempst P, Kornberg RD (2002) A complex of the Srb8, -9, -10, and -11 transcriptional regulatory proteins from yeast. *J Biol Chem* **277**: 44202-44207.

Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV (2002) The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* **55**: 244-265.

Bouazzi H, Lesca G, Trujillo C, Alwasiyah MK, Munnich A (2015) Nonsyndromic X-linked intellectual deficiency in three brothers with a novel MED12 missense mutation [c.5922G>T (p.Glu1974His). *Clin Case Rep* **3**: 604-609.

Bourbon HM (2008) Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res* **36**: 3993-4008.

Bourbon HM, Aguilera A, Ansari AZ, Asturias FJ, Berk AJ, Bjorklund S, Blackwell TK, Borggrefe T, Carey M, Carlson M, Conaway JW, Conaway RC, Emmons SW, Fondell JD, Freedman LP, Fukasawa T, Gustafsson CM, Han M, He X, Herman PK, Hinnebusch AG, Holmberg S, Holstege FC, Jaehning JA, Kim YJ, Kuras L, Leutz A, Lis JT, Meisterernest M, Näär AM, Nasmyth K, Parvin JD, Ptashne M, Reinberg D, Ronne H, Sadowski I, Sakurai H, Sipiczki M, Sternberg PW, Stillman DJ, Strich R, Struhl K, Svejstrup JQ, Tuck S, Winston F, Roeder RG, Kornberg RD (2004) A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. *Mol Cell* **14**: 553-557.

Bourgeron T, Chretien D, Poggi-Bach J, Doonan S, Rabier D, Letouze P, Munnich A, Rotig A, Landrieu P, Rustin P (1994) Mutation of the fumarase gene in two siblings with progressive encephalopathy and fumarase deficiency. *J Clin Invest* **93**: 2514-2518.

Brooks SE, Zhan M, Cote T, Baquet CR (2004) Surveillance, epidemiology, and end results analysis of 2677 cases of uterine sarcoma 1989-1999. *Gynecol Oncol* **93**: 204-208.

Brunen D, Willems SM, Kellner U, Midgley R, Simon I, Bernards R (2013) TGF-beta: an emerging player in drug resistance. *Cell Cycle* **12**: 2960-2968.

Burns MB, Temiz NA, Harris RS (2013) Evidence for APOBEC3B mutagenesis in multiple human cancers. *Nat Genet* **45**: 977-983.

Cancer Genome Atlas Network (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**: 330-337.

Cancer Genome Atlas Research Network (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* **474**: 609-615.

Canevari RA, Pontes A, Rosa FE, Rainho CA, Rogatto SR (2005) Independent clonal origin of multiple uterine leiomyomas that was determined by X chromosome inactivation and microsatellite analysis. *Am J Obstet Gynecol* **193**: 1395-1403.

Cani AK, Hovelson DH, McDaniel AS, Sadis S, Haller MJ, Yadati V, Amin AM, Bratley J, Bandla S, Williams PD, Rhodes K, Liu CJ, Quist MJ, Rhodes DR, Grasso CS, Kleer CG, Tomlins SA (2015) Next-Gen Sequencing Exposes Frequent MED12 Mutations and Actionable Therapeutic Targets in Phyllodes Tumors. *Mol Cancer Res* **4**: 613-619.

Cardozo ER, Clark AD, Banks NK, Henne MB, Stegmann BJ, Segars JH (2012) The estimated annual cost of uterine leiomyomata in the United States. *Am J Obstet Gynecol* **206**: 211.e1-211.e9.

Chen L, Widhopf G, Huynh L, Rassenti L, Rai KR, Weiss A, Kipps TJ (2002) Expression of ZAP-70 is associated with increased B-cell receptor signaling in chronic lymphocytic leukemia. *Blood* **100**: 4609-4614.

Choi JD and Lee JS (2013) Interplay between Epigenetics and Genetics in Cancer. *Genomics Inform* **11**: 164-173.

Christacos NC, Quade BJ, Dal Cin P, Morton CC (2006) Uterine leiomyomata with deletions of Ip represent a distinct cytogenetic subgroup associated with unusual histologic features. *Genes Chromosomes Cancer* **45**: 304-312.

Clark AD, Oldenbroek M, Boyer TG (2015) Mediator kinase module and human tumorigenesis. *Crit Rev Biochem Mol Biol* **50**: 393-426.

Claus R, Lucas DM, Ruppert AS, Williams KE, Weng D, Patterson K, Zucknick M, Oakes CC, Rassenti LZ, Greaves AW, Geyer S, Wierda WG, Brown JR, Gribben JG, Barrientos JC, Rai KR, Kay NE, Kipps TJ, Shields P, Zhao W, Grever MR, Plass C, Byrd JC (2014) Validation of ZAP-70 methylation and its relative significance in predicting outcome in chronic lymphocytic leukemia. *Blood* **124**: 42-48.

Cramer SF and Patel A (1990) The frequency of uterine leiomyomas. *Am J Clin Pathol* **94**: 435-438.

Croce CM (2008) Oncogenes and cancer. *N Engl J Med* **358**: 502-511.

Dal Cin P, Vanni R, Marras S, Moerman P, Kools P, Andria M, Valdes E, Deprest J, Van de Ven W, Van den Berghe H (1995) Four cytogenetic subgroups can be identified in endometrial polyps. *Cancer Res* **55**: 1565-1568.

Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, Buchbinder A, Budman D, Dittmar K, Kolitz J, Lichtman SM, Schulman P, Vinciguerra VP, Rai KR, Ferrarini M, Chiorazzi N (1999) Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* **94**: 1840-1847.

Damm F, Mylonas E, Cosson A, Yoshida K, Della Valle V, Mouly E, Diop M, Scourzic L, Shiraishi Y, Chiba K, Tanaka H, Miyano S, Kikushige Y, Davi F, Lambert J, Gautheret D, Merle-Beral H, Sutton L, Dessen P, Solary E, Akashi K, Vainchenker W, Mercher T, Droin N, Ogawa S, Nguyen-Khac F, Bernard OA (2014) Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov* **4**: 1088-1101.

Dave UP, Akagi K, Tripathi R, Cleveland SM, Thompson MA, Yi M, Stephens R, Downing JR, Jenkins NA, Copeland NG (2009) Murine leukemias with retroviral insertions at Lmo2 are predictive of the leukemias induced in SCID-X1 patients following retroviral gene therapy. *PLoS Genet* **5**: e1000491.

Day Baird D, Dunson DB, Hill MC, Cousins D, Schectman JM (2003) High cumulative incidence of uterine leiomyoma in black and white women: ultrasound evidence. *Am J Obstet Gynecol* **188**: 100-107.

de Graaff MA, Cleton-Jansen AM, Szuhai K, Bovee JV (2013) Mediator complex subunit 12 exon 2 mutation analysis in different subtypes of smooth muscle tumors confirms genetic heterogeneity. *Hum Pathol* **44**: 1597-1604.

Diederichs S, Bartsch L, Berkmann JC, Frose K, Heitmann J, Hoppe C, Iggena D, Jazmati D, Karschnia P, Linsenmeier M, Maulhardt T, Mohrmann L, Morstein J, Paffenholz SV, Ropenack P, Ruckert T, Sandig L, Schell M, Steinmann A, Voss G, Wasmuth J, Weinberger ME, Wullenkord R (2016) The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. *EMBO Mol Med* **8**: 442-457.

Ding N, Tomomori-Sato C, Sato S, Conaway RC, Conaway JW, Boyer TG (2009) MED19 and MED26 are synergistic functional targets of the RE1 silencing transcription factor in epigenetic silencing of neuronal gene expression. *J Biol Chem* **284**: 2648-2656.

Ding N, Zhou H, Esteve PO, Chin HG, Kim S, Xu X, Joseph SM, Friez MJ, Schwartz CE, Pradhan S, Boyer TG (2008) Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. *Mol Cell* **31**: 347-359.

Donner AJ, Szostek S, Hoover JM, Espinosa JM (2007) CDK8 is a stimulus-specific positive coregulator of p53 target genes. *Mol Cell* **27**: 121-133.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-1797.

Elmlund H, Baraznenok V, Lindahl M, Samuelsen CO, Koeck PJ, Holmberg S, Hebert H, Gustafsson CM (2006) The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II. *Proc Natl Acad Sci U S A* **103**: 15788-15793.

Fabbri G and Dalla-Favera R (2016) The molecular pathogenesis of chronic lymphocytic leukaemia. *Nat Rev Cancer* **16**: 145-162.

Faerstein E, Szklo M, Rosenshein N (2001) Risk factors for uterine leiomyoma: a practice-based case-control study. I. African-American heritage, reproductive history, body size, and smoking. *Am J Epidemiol* **153**: 1-10.

Fais F, Ghiotto F, Hashimoto S, Sellars B, Valetto A, Allen SL, Schulman P, Vinciguerra VP, Rai K, Rassenti LZ, Kipps TJ, Dighiero G, Schroeder HW,Jr, Ferrarini M, Chiorazzi N (1998) Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J Clin Invest* **102**: 1515-1525.

Fearon ER and Vogelstein B (1990) A genetic model for colorectal tumorigenesis. *Cell* **61**: 759-767.

Feinberg AP, Koldobskiy MA, Gondor A (2016) Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat Rev Genet* **17**: 284-299.

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* **44**: D279-85.

Firestein R, Bass AJ, Kim SY, Dunn IF, Silver SJ, Guney I, Freed E, Ligon AH, Vena N, Ogino S, Chheda MG, Tamayo P, Finn S, Shrestha Y, Boehm JS, Jain S, Bojarski E, Mermel C, Barretina J, Chan JA, Baselga J, Tabernero J, Root DE, Fuchs CS, Loda M, Shivdasani RA, Meyerson M, Hahn WC (2008) CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. *Nature* **455**: 547-551.

Flake GP, Andersen J, Dixon D (2003) Etiology and pathogenesis of uterine leiomyomas: a review. *Environ Health Perspect* **111**: 1037-1054.

Forbes SA, Bhamra G, Bamford S, Dawson E, Kok C, Clements J, Menzies A, Teague JW, Futreal PA, Stratton MR (2008) The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr Protoc Hum Genet* **Chapter 10**: Unit 10.11.

Forment JV, Kaidi A, Jackson SP (2012) Chromothripsis and cancer: causes and consequences of chromosome shattering. *Nat Rev Cancer* **12**: 663-670.

Frizzell N, Lima M, Baynes JW (2011) Succination of proteins in diabetes. *Free Radic Res* **45**: 101-109.

Fryns JP and Buttiens M (1987) X-linked mental retardation with marfanoid habitus. *Am J Med Genet* **28**: 267-274.

Fusco A and Fedele M (2007) Roles of HMGA proteins in cancer. *Nat Rev Cancer* **7**: 899-910.

Gadducci A, Landoni F, Sartori E, Zola P, Maggino T, Lissoni A, Bazzurini L, Arisio R, Romagnolo C, Cristofani R (1996) Uterine leiomyosarcoma: analysis of treatment failures and survival. *Gynecol Oncol* **62**: 25-32.

Galbraith MD, Allen MA, Bensard CL, Wang X, Schwinn MK, Qin B, Long HW, Daniels DL, Hahn WC, Dowell RD, Espinosa JM (2013) HIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. *Cell* **153**: 1327-1339.

Galbraith MD, Donner AJ, Espinosa JM (2010) CDK8: a positive regulator of transcription. *Transcription* **1**: 4-12.

Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C, Schultz N (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**: pl1.

Garcia-Torres R, Cruz D, Orozco L, Heidet L, Gubler MC (2000) Alport syndrome and diffuse leiomyomatosis. Clinical aspects, pathology, molecular biology and extracellular matrix studies. A synthesis. *Nephrologie* **21**: 9-12.

Gardie B, Remenieras A, Kattygnarath D, Bombled J, Lefevre S, Perrier-Trudova V, Rustin P, Barrois M, Slama A, Avril MF, Bessis D, Caron O, Caux F, Collignon P, Coupier I, Cremin C, Dollfus H, Dugast C, Escudier B, Faivre L, Field M, Gilbert-Dussardier B, Janin N, Leport Y, Leroux D, Lipsker D, Malthieu F, McGilliwray B, Maugard C, Mejean A, Mortemousque I, Plessis G, Poppe B, Pruvost-Balland C, Rooker S, Roume J, Soufir N, Steinraths M, Tan MH, Theodore C, Thomas L, Vabres P, Van Glabeke E, Meric JB, Verkarre V, Lenoir G, Joulin V, Deveaux S, Cusin V, Feunteun J, Teh BT, Bressac-de Paillerets B, Richard S, French National Cancer Institute "Inherited predisposition to kidney cancer" network (2011) Novel FH mutations in families with hereditary leiomyomatosis and renal cell cancer (HLRCC) and patients with isolated type 2 papillary renal cell carcinoma. *J Med Genet* **48**: 226-234.

Gartler SM and Goldman MA (2001) Biology of the X chromosome. *Curr Opin Pediatr* **13**: 340-345.

Gattas GJ, Quade BJ, Nowak RA, Morton CC (1999) HMGIC expression in human adult and fetal tissues and in uterine leiomyomata. *Genes Chromosomes Cancer* **25**: 316-322.

Giuntoli RL,2nd, Metzinger DS, DiMarco CS, Cha SS, Sloan JA, Keeney GL, Gostout BS (2003) Retrospective review of 208 patients with leiomyosarcoma of the uterus: prognostic indicators, surgical management, and adjuvant therapy. *Gynecol Oncol* **89**: 460-469.

Gobert V, Osman D, Bras S, Auge B, Boube M, Bourbon HM, Horn T, Boutros M, Haenlin M, Waltzer L (2010) A genome-wide RNA interference screen identifies a differential role of the mediator CDK8 module subunits for GATA/ RUNX-activated transcription in Drosophila. *Mol Cell Biol* **30**: 2837-2848.

Graham JM,Jr and Schwartz CE (2013) MED12 related disorders. *Am J Med Genet A* **161A**: 2734-2740.

Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, Vastrik I, Schmidt EE, Avis T, Barthorpe S, Bhamra G, Buck G, Choudhury B, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Cahill DP, Louis DN, Goldstraw P, Nicholson AG, Brasseur F, Looijenga L, Weber BL, Chiew YE, DeFazio A, Greaves MF, Green AR, Campbell P, Birney E, Easton DF, Chenevix-Trench G, Tan MH, Khoo SK, Teh BT, Yuen ST, Leung SY, Wooster R, Futreal PA, Stratton MR (2007) Patterns of somatic mutation in human cancer genomes. *Nature* **446**: 153-158.

Grubb RL,3rd, Franks ME, Toro J, Middelton L, Choyke L, Fowler S, Torres-Cabala C, Glenn GM, Choyke P, Merino MJ, Zbar B, Pinto PA, Srinivasan R, Coleman JA, Linehan WM (2007) Hereditary leiomyomatosis and renal cell cancer: a syndrome associated with an aggressive form of inherited renal cancer. *J Urol* **177**: 2074-9; discussion 2079-80.

Guieze R, Robbe P, Clifford R, de Guibert S, Pereira B, Timbs A, Dilhuydy MS, Cabes M, Ysebaert L, Burns A, Nguyen-Khac F, Davi F, Veronese L, Combes P, Le Garff-Tavernier M, Leblond V, Merle-Beral H, Alsolami R, Hamblin A, Mason J, Pettitt A, Hillmen P, Taylor J, Knight SJ, Tournilhac O, Schuh A (2015) Presence of multiple recurrent mutations confers poor trial outcome of relapsed/refractory CLL. *Blood* **126**: 2110-2117.

Gutierrez A,Jr, Tschumper RC, Wu X, Shanafelt TD, Eckel-Passow J, Huddleston PM,3rd, Slager SL, Kay NE, Jelinek DF (2010) LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. *Blood* **116**: 2975-2983.

Halder SK, Laknaur A, Miller J, Layman LC, Diamond M, Al-Hendy A (2015) Novel MED12 gene somatic mutations in women from the Southern United States with symptomatic uterine fibroids. *Mol Genet Genomics* **290**: 505-511.

Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK (1999) Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* **94**: 1848-1854.

Hanahan D and Coussens LM (2012) Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* **21**: 309-322.

Hanahan D and Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* **144**: 646-674.

Hanahan D and Weinberg RA (2000) The hallmarks of cancer. *Cell* **100**: 57-70.

Harrison WJ, Andrici J, Maclean F, Madadi-Ghahan R, Farzin M, Sioson L, Toon CW, Clarkson A, Watson N, Pickett J, Field M, Crook A, Tucker K, Goodwin A, Anderson L, Srinivasan B, Grossmann P, Martinek P, Ondic O, Hes O, Trpkov K, Clifton-Bligh RJ, Dwight T, Gill AJ (2015) Fumarate Hydratase-deficient Uterine Leiomyomas Occur in Both the Syndromic and Sporadic Settings. *Am J Surg Pathol*.

Heinonen HR, Sarvilinna NS, Sjöberg J, Kämpjärvi K, Pitkänen E, Vahteristo P, Mäkinen N, Aaltonen LA (2014) MED12 mutation frequency in unselected sporadic uterine leiomyomas. *Fertil Steril* **102**: 1137-1142.

Hengartner CJ, Myer VE, Liao SM, Wilson CJ, Koh SS, Young RA (1998) Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol Cell* **2**: 43-53.

Hodge JC, Kim TM, Dreyfuss JM, Somasundaram P, Christacos NC, Rousselle M, Quade BJ, Park PJ, Stewart EA, Morton CC (2012) Expression profiling of uterine leiomyomata cytogenetic subgroups reveals distinct signatures in matched myometrium: transcriptional profilingof the t(12;14) and evidence in support of predisposing genetic heterogeneity. *Hum Mol Genet* **21**: 2312-2329.

Hodges KB, Abdul-Karim FW, Wang M, Lopez-Beltran A, Montironi R, Easley S, Zhang S, Wang N, MacLennan GT, Cheng L (2009) Evidence for transformation of fibroadenoma of the breast to malignant phyllodes tumor. *Appl Immunohistochem Mol Morphol* **17**: 345-350.

Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A, Kadel S, Moll I, Nagore E, Hemminki K, Schadendorf D, Kumar R (2013) TERT promoter mutations in familial and sporadic melanoma. *Science* **339**: 959-961.

Hosking FJ, Dobbins SE, Houlston RS (2011) Genome-wide association studies for detecting cancer susceptibility. *Br Med Bull* **97**: 27-46.

Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, De Die-Smulders C, Persky N, Grompe M, Joenje H, Pals G, Ikeda H, Fox EA, D'Andrea AD (2002) Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* **297**: 606-609.

Huang S, Holzel M, Knijnenburg T, Schlicker A, Roepman P, McDermott U, Garnett M, Grernrum W, Sun C, Prahallad A, Groenendijk FH, Mittempergher L, Nijkamp W, Neefjes J, Salazar R, Ten Dijke P, Uramoto H, Tanaka F, Beijersbergen RL, Wessels LF, Bernards R (2012) MED12 controls the response to multiple cancer drugs through regulation of TGF-beta receptor signaling. *Cell* **151**: 937-950.

Hulea L and Nepveu A (2012) CUX1 transcription factors: from biochemical activities and cell-based assays to mouse models and human diseases. *Gene* **497**: 18-26.

Ingraham SE, Lynch RA, Kathiresan S, Buckler AJ, Menon AG (1999) hREC2, a RAD51-like gene, is disrupted by t(12;14) (q15;q24.1) in a uterine leiomyoma. *Cancer Genet Cytogenet* **115**: 56-61.

Isaacs JS, Jung YJ, Mole DR, Lee S, Torres-Cabala C, Chung YL, Merino M, Trepel J, Zbar B, Toro J, Ratcliffe PJ, Linehan WM, Neckers L (2005) HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: novel role of fumarate in regulation of HIF stability. *Cancer Cell* **8**: 143-153.

Isidor B, Lefebvre T, Le Vaillant C, Caillaud G, Faivre L, Jossic F, Joubert M, Winer N, Le Caignec C, Borck G, Pelet A, Amiel J, Toutain A, Ronce N, Raynaud M, Verloes A, David A (2014) Blepharophimosis, short humeri, developmental delay and hirschsprung disease: expanding the phenotypic spectrum of MED12 mutations. *Am J Med Genet A* **164A**: 1821-1825.

Ito M, Yuan CX, Malik S, Gu W, Fondell JD, Yamamura S, Fu ZY, Zhang X, Qin J, Roeder RG (1999) Identity between TRAP and SMCC complexes indicates novel pathways for the function of nuclear receptors and diverse mammalian activators. *Mol Cell* **3**: 361-370.

Je EM, Kim MR, Min KO, Yoo NJ, Lee SH (2012) Mutational analysis of MED12 exon 2 in uterine leiomyoma and other common tumors. *Int J Cancer* **131**: E1044-E1047.

Joensuu T, Hämäläinen R, Yuan B, Johnson C, Tegelberg S, Gasparini P, Zelante L, Pirvola U, Pakarinen L, Lehesjoki AE, de la Chapelle A, Sankila EM (2001) Mutations in a novel gene with transmembrane domains underlie Usher syndrome type 3. *Am J Hum Genet* **69**: 673-684.

Jori B, Kamps R, Xanthoulea S, Delvoux B, Blok MJ, Van de Vijver KK, de Koning B, Oei FT, Tops CM, Speel EJ, Kruitwagen RF, Gomez-Garcia EB, Romano A (2015) Germ-line variants identified by next generation sequencing in a panel of estrogen and cancer associated genes correlate with poor clinical outcome in Lynch syndrome patients. *Oncotarget* **6**: 41108-41122.

Joseph NM, Solomon DA, Frizzell N, Rabban JT, Zaloudek C, Garg K (2015) Morphology and Immunohistochemistry for 2SC and FH Aid in Detection of Fumarate Hydratase Gene Aberrations in Uterine Leiomyomas From Young Patients. *Am J Surg Pathol* **39**: 1529-1539.

Jung H, Lee D, Lee J, Park D, Kim YJ, Park WY, Hong D, Park PJ, Lee E (2015) Intron retention is a widespread mechanism of tumor-suppressor inactivation. *Nat Genet* **47**: 1242-1248.

Junttila MR and de Sauvage FJ (2013) Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* **501**: 346-354.

Kaucka M, Plevova K, Pavlova S, Janovska P, Mishra A, Verner J, Prochazkova J, Krejci P, Kotaskova J, Ovesna P, Tichy B, Brychtova Y, Doubek M, Kozubik A, Mayer J, Pospisilova S, Bryja V (2013) The planar cell polarity pathway drives pathogenesis of chronic lymphocytic leukemia by the regulation of B-lymphocyte migration. *Cancer Res* **73**: 1491-1501.

Keightley MC, Layton JE, Hayman JW, Heath JK, Lieschke GJ (2011) Mediator subunit 12 is required for neutrophil development in zebrafish. *PLoS One* **6**: e23845.

Khurana E, Fu Y, Chakravarty D, Demichelis F, Rubin MA, Gerstein M (2016) Role of non-coding sequence variants in cancer. *Nat Rev Genet* **17**: 93-108.

Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Bjorkman M, Mpindi JP, Haapa-Paananen S, Vainio P, Edgren H, Wolf M, Astola J, Nees M, Hautaniemi S, Kallioniemi O (2008) Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9**: R139.

Kim S, Xu X, Hecht A, Boyer TG (2006) Mediator is a transducer of Wnt/beta-catenin signaling. *J Biol Chem* **281**: 14066-14075.

Kinzler KW and Vogelstein B (1998) Landscaping the cancer terrain. *Science* **280**: 1036-1037.

Kinzler KW and Vogelstein B (1997) Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature* **386**: 761, 763.

Kipps TJ (2007) The B-cell receptor and ZAP-70 in chronic lymphocytic leukemia. *Best Pract Res Clin Haematol* **20**: 415-424.

Kitano T, Schwarz C, Nickel B, Paabo S (2003) Gene diversity patterns at 10 X-chromosomal loci in humans and chimpanzees. *Mol Biol Evol* **20**: 1281-1289.

Kiuru M, Launonen V, Hietala M, Aittomäki K, Vierimaa O, Salovaara R, Arola J, Pukkala E, Sistonen P, Herva R, Aaltonen LA (2001) Familial cutaneous leiomyomatosis is a two-hit condition associated with renal cell cancer of characteristic histopathology. *Am J Pathol* **159**: 825-829.

Kiuru M, Lehtonen R, Arola J, Salovaara R, Järvinen H, Aittomäki K, Sjöberg J, Visakorpi T, Knuutila S, Isola J, Delahunt B, Herva R, Launonen V, Karhu A, Aaltonen LA (2002) Few FH mutations in sporadic counterparts of tumor types observed in hereditary leiomyomatosis and renal cell cancer families. *Cancer Res* **62**: 4554-4557.

Kjerulff KH, Langenberg P, Seidman JD, Stolley PD, Guzinski GM (1996) Uterine leiomyomas. Racial differences in severity, symptoms and age at diagnosis. *J Reprod Med* **41**: 483-490.

Klemke M, Meyer A, Nezhad MH, Bartnitzke S, Drieschner N, Frantzen C, Schmidt EH, Belge G, Bullerdiek J (2009) Overexpression of HMGA2 in uterine leiomyomas points to its general role for the pathogenesis of the disease. *Genes Chromosomes Cancer* **48**: 171-178.

Kloosterman WP, Koster J, MoleNäär JJ (2014) Prevalence and clinical implications of chromothripsis in cancer genomes. *Curr Opin Oncol* **26**: 64-72.

Knudson AG,Jr (1971) Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A* **68**: 820-823.

Knuesel MT, Meyer KD, Bernecky C, Taatjes DJ (2009a) The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev* **23**: 439-451.

Knuesel MT, Meyer KD, Donner AJ, Espinosa JM, Taatjes DJ (2009b) The human CDK8 subcomplex is a histone kinase that requires Med12 for activity and can function independently of mediator. *Mol Cell Biol* **29**: 650-661.

Koivisto-Korander R, Martinsen JI, Weiderpass E, Leminen A, Pukkala E (2012) Incidence of uterine leiomyosarcoma and endometrial stromal sarcoma in Nordic countries: results from NORDCAN and NOCCA databases. *Maturitas* **72**: 56-60.

Koivunen P, Hirsilä M, Remes AM, Hassinen IE, Kivirikko KI, Myllyharju J (2007) Inhibition of hypoxia-inducible factor (HIF) hydroxylases by citric acid cycle intermediates: possible links between cell metabolism and stabilization of HIF. *J Biol Chem* **282**: 4524-4532.

Kontro M, Kuusanmäki H, Eldfors S, Burmeister T, Andersson EI, Bruserud O, Brummendorf TH, Edgren H, Gjertsen BT, Itala-Remes M, Lagstrom S, Lohi O, Lundan T, Marti JM, Majumder MM, Parsons A, Pemovska T, Rajala H, Vettenranta K, Kallioniemi O, Mustjoki S, Porkka K, Heckman CA (2014) Novel activating STAT5B mutations as putative drivers of T-cell acute lymphoblastic leukemia. *Leukemia* **28**: 1738-1742.

Koski TA (2010) Molecular genetic background of tumours in hereditary leiomyomatosis and renal cell cancer syndrome. University of Helsinki.

Kosugi S, Hasebe M, Tomita M, Yanagawa H (2009) Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A* **106**: 10171-10176.

Kuijper A, Buerger H, Simon R, Schaefer KL, Croonen A, Boecker W, van der Wall E, van Diest PJ (2002) Analysis of the progression of fibroepithelial tumours of the breast by PCR-based clonality assay. *J Pathol* **197**: 575-581.

Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* **4**: 1073-1081.

Kurosaki T and Maquat LE (2016) Nonsense-mediated mRNA decay in humans at a glance. *J Cell Sci* **129**: 461-467.

Kämpjärvi K, Kim NH, Keskitalo S, Clark AD, von Nandelstadh P, Turunen M, Heikkinen T, Park MJ, Mäkinen N, Kivinummi K, Lintula S, Hotakainen K, Nevanlinna H, Hokland P, Böhling T, Bützow R, Böhm J, Mecklin JP, Järvinen H, Kontro M, Visakorpi T, Taipale J, Varjosalo M, Boyer TG, Vahteristo P (2016) Somatic MED12 mutations in prostate cancer and uterine leiomyomas promote tumorigenesis through distinct mechanisms. *Prostate* **76**: 22-31.

Lahiri DK and Nurnberger JI,Jr (1991) A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. *Nucleic Acids Res* **19**: 5444.

Landau DA, Carter SL, Stojanov P, McKenna A, Stevenson K, Lawrence MS, Sougnez C, Stewart C, Sivachenko A, Wang L, Wan Y, Zhang W, Shukla SA, Vartanov A, Fernandes SM, Saksena G, Cibulskis K, Tesar B, Gabriel S, Hacohen N, Meyerson M, Lander ES, Neuberg D, Brown JR, Getz G, Wu CJ (2013) Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**: 714-726.

Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, Kluth S, Bozic I, Lawrence M, Bottcher S, Carter SL, Cibulskis K, Mertens D, Sougnez CL, Rosenberg M, Hess JM, Edelmann J, Kless S, Kneba M, Ritgen M, Fink A, Fischer K, Gabriel S, Lander ES, Nowak MA, Dohner H, Hallek M, Neuberg D, Getz G, Stilgenbauer S, Wu CJ (2015) Mutations driving CLL and their evolution in progression and relapse. *Nature* **526**: 525-530.

Langley KG, Brown J, Gerber RJ, Fox J, Friez MJ, Lyons M, Schrier Vergano SA (2015) Beyond Ohdo syndrome: A familial missense mutation broadens the MED12 spectrum. *Am J Med Genet A* **167**: 3180-3185.

Launonen V, Vierimaa O, Kiuru M, Isola J, Roth S, Pukkala E, Sistonen P, Herva R, Aaltonen LA (2001) Inherited susceptibility to uterine leiomyomas and renal cell cancer. *Proc Natl Acad Sci U S A* **98**: 3387-3392.

Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, Drier Y, Zou L, Ramos AH,

Pugh TJ, Stransky N, Helman E, Kim J, Sougnez C, Ambrogio L, Nickerson E, Shefler E, Cortes ML, Auclair D, Saksena G, Voet D, Noble M, DiCara D, Lin P, Lichtenstein L, Heiman DI, Fennell T, Imielinski M, Hernandez B, Hodis E, Baca S, Dulak AM, Lohr J, Landau DA, Wu CJ, Melendez-Zajgla J, Hidalgo-Miranda A, Koren A, McCarroll SA, Mora J, Lee RS, Crompton B, Onofrio R, Parkin M, Winckler W, Ardlie K, Gabriel SB, Roberts CW, Biegel JA, Stegmaier K, Bass AJ, Garraway LA, Meyerson M, Golub TR, Gordenin DA, Sunyaev S, Lander ES, Getz G (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**: 214-218.

Lee RS, Stewart C, Carter SL, Ambrogio L, Cibulskis K, Sougnez C, Lawrence MS, Auclair D, Mora J, Golub TR, Biegel JA, Getz G, Roberts CW (2012) A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. *J Clin Invest* **122**: 2983-2988.

Lehtonen HJ (2011) Hereditary leiomyomatosis and renal cell cancer: update on clinical and molecular characteristics. *Fam Cancer* **10**: 397-411.

Lehtonen HJ, Kiuru M, Ylisaukko-Oja SK, Salovaara R, Herva R, Koivisto PA, Vierimaa O, Aittomäki K, Pukkala E, Launonen V, Aaltonen LA (2006) Increased risk of cancer in patients with fumarate hydratase germline mutation. *J Med Genet* **43**: 523-526.

Lehtonen R, Kiuru M, Vanharanta S, Sjöberg J, Aaltonen LM, Aittomäki K, Arola J, Bützow R, Eng C, Husgafvel-Pursiainen K, Isola J, Järvinen H, Koivisto P, Mecklin JP, Peltomäki P, Salovaara R, Wasenius VM, Karhu A, Launonen V, Nupponen NN, Aaltonen LA (2004) Biallelic inactivation of fumarate hydratase (FH) occurs in nonsyndromic uterine leiomyomas but is rare in other tumors. *Am J Pathol* **164**: 17-22.

Leibsohn S, d'Ablaing G, Mishell DR,Jr, Schlaerth JB (1990) Leiomyosarcoma in a series of hysterectomies performed for presumed uterine leiomyomas. *Am J Obstet Gynecol* **162**: 968-74; discussion 974-6.

Lesca G, Moizard MP, Bussy G, Boggio D, Hu H, Haas SA, Ropers HH, Kalscheuer VM, Des Portes V, Labalme A, Sanlaville D, Edery P, Raynaud M, Lespinasse J (2013) Clinical and neurocognitive characterization of a family with a novel MED12 gene frameshift mutation. *Am J Med Genet A* **161A**: 3063-3071.

Levy-Lahad E and Friedman E (2007) Cancer risks among BRCA1 and BRCA2 mutation carriers. *Br J Cancer* **96**: 11-15.

Li N, Fassl A, Chick J, Inuzuka H, Li X, Mansour MR, Liu L, Wang H, King B, Shaik S, Gutierrez A, Ordureau A, Otto T, Kreslavsky T, Baitsch L, Bury L, Meyer CA, Ke N, Mulry KA, Kluk MJ, Roy M, Kim S, Zhang X, Geng Y, Zagozdzon A, Jenkinson S, Gale RE, Linch DC, Zhao JJ, Mullighan CG, Harper JW, Aster JC, Aifantis I, von Boehmer H, Gygi SP, Wei W, Look AT, Sicinski P (2014) Cyclin C is a haploinsufficient tumour suppressor. *Nat Cell Biol* **16**: 1080-1091.

Liegl-Atzwanger B, Heitzer E, Flicker K, Muller S, Ulz P, Saglam O, Tavassoli F, Devouassoux-Shisheboran M, Geigl J, Moinfar F (2016) Exploring chromosomal abnormalities and genetic changes in uterine smooth muscle tumors. *Mod Pathol* Epub ahead of print.

Lien HC, Huang CS, Yang YW, Jeng YM (2016) Mutational analysis of MED12 exon 2 in a spectrum of fibroepithelial tumours of the breast: implications for pathogenesis and histogenesis. *Histopathology* **68**: 433-441.

Ligon AH and Morton CC (2000) Genetics of uterine leiomyomata. *Genes Chromosomes Cancer* **28**: 235-245.

Lim WK, Ong CK, Tan J, Thike AA, Ng CC, Rajasegaran V, Myint SS, Nagarajan S, Nasir ND, McPherson JR, Cutcutache I, Poore G, Tay ST, Ooi WS, Tan VK, Hartman M, Ong KW, Tan BK, Rozen SG, Tan PH, Tan P, Teh BT (2014) Exome sequencing identifies highly recurrent MED12 somatic mutations in breast fibroadenoma. *Nat Genet* **46**: 877-880.

Lin JR and Hu J (2013) SeqNLS: nuclear localization signal prediction based on frequent pattern mining and linear motif scoring. *PLoS One* **8**: e76864.

Lindeboom RG, Supek F, Lehner B (2016) The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat Genet* **48**: 1112-1118.

Linder D and Gartler SM (1965) Glucose-6-phosphate dehydrogenase mosaicism: utilization as a cell marker in the study of leiomyomas. *Science* **150**: 67-69.

Liu Y, Easton J, Shao Y, Wilkinson M, Edmonson M, Ma X, Smith M, Rusch M, Jaime Guidry Auvil J, Gerhard D, Relling M, Winick N, Raetz E, Devidas M, Willman C, Harvey R, Carroll W, Dunsmore K, Winter S, Wood B, Downing J, Loh M, Hunger S, Zhang J, Mullighan C (2015) The Genomic Landscape of Childhood T-Lineage Acute Lymphoblastic Leukemia. 126: 691.

Ljungstrom V, Cortese D, Young E, Pandzic T, Mansouri L, Plevova K, Ntoufa S, Baliakas P, Clifford R, Sutton LA, Blakemore SJ, Stavroyianni N, Agathangelidis A, Rossi D, Hoglund M, Kotaskova J, Juliusson G, Belessi C, Chiorazzi N, Panagiotidis P, Langerak AW, Smedby KE, Oscier D, Gaidano G, Schuh A, Davi F, Pott C, Strefford JC, Trentin L, Pospisilova S, Ghia P, Stamatopoulos K, Sjoblom T, Rosenquist R (2016) Whole-exome sequencing in relapsing chronic lymphocytic leukemia: clinical impact of recurrent RPS15 mutations. *Blood* **127**: 1007-1016.

Lu D, Zhao Y, Tawatao R, Cottam HB, Sen M, Leoni LM, Kipps TJ, Corr M, Carson DA (2004) Activation of the Wnt signaling pathway in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* **101**: 3118-3123.

Lujan JE, Carlin ME, Lubs HA (1984) A form of X-linked mental retardation with marfanoid habitus. *Am J Med Genet* **17**: 311-322.

Lyon MF (1961) Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* **190**: 372-373.

Maat-Kievit A, Brunner HG, Maaswinkel-Mooij P (1993) Two additional cases of the Ohdo blepharophimosis syndrome. *Am J Med Genet* **47**: 901-906.

Major FJ, Blessing JA, Silverberg SG, Morrow CP, Creasman WT, Currie JL, Yordan E, Brady MF (1993) Prognostic factors in early-stage uterine sarcoma. A Gynecologic Oncology Group study. *Cancer* **71**: 1702-1709.

Malik S and Roeder RG (2010) The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat Rev Genet* **11**: 761-772.

Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. *Science* **298**: 1912-1934.

Margolskee E, Jobanputra V, Jain P, Chen J, Ganapathi K, Nahum O, Levy B, Morscio J, Murty V, Tousseyn T, Alobeid B, Mansukhani M, Bhagat G (2016) Genetic landscape of T- and NK-cell post-transplant lymphoproliferative disorders. *Oncotarget* Epub ahead of print.

Markowski DN, Bartnitzke S, Loning T, Drieschner N, Helmke BM, Bullerdiek J (2012) MED12 mutations in uterine fibroids-their relationship to cytogenetic subgroups. *Int J Cancer*: 1528-1536.

Markowski DN, Helmke BM, Bartnitzke S, Loning T, Bullerdiek J (2014) Uterine fibroids: do we deal with more than one disease? *Int J Gynecol Pathol* **33**: 568-572.

Markowski DN, Holzmann C, Bullerdiek J (2015) Genetic alterations in uterine fibroids - a new direction for pharmacological intervention? *Expert Opin Ther Targets* **19**: 1485-1494.

Markowski DN, Huhle S, Nimzyk R, Stenman G, Loning T, Bullerdiek J (2013) MED12 mutations occurring in benign and malignant mammalian smooth muscle tumors. *Genes Chromosomes Cancer*: 297-304.

Marshall LM, Spiegelman D, Goldman MB, Manson JE, Colditz GA, Barbieri RL, Stampfer MJ, Hunter DJ (1998) A prospective study of reproductive factors and oral contraceptive use in relation to the risk of uterine leiomyomata. *Fertil Steril* **70**: 432-439.

Martincorena I and Campbell PJ (2015) Somatic mutation in cancer and normal cells. *Science* **349**: 1483-1489.

Matsubara A, Sekine S, Yoshida M, Yoshida A, Taniguchi H, Kushima R, Tsuda H, Kanai Y (2013) Prevalence of MED12 mutations in uterine and extrauterine smooth muscle tumours. *Histopathology* **62**: 657-661.

Mayr C, Hemann MT, Bartel DP (2007) Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. *Science* **315**: 1576-1579.

McGranahan N and Swanton C (2015) Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell* **27**: 15-26.

McGuire MM, Yatsenko A, Hoffner L, Jones M, Surti U, Rajkovic A (2012) Whole Exome Sequencing in a Random Sample of North American Women with Leiomyomas Identifies MED12 Mutations in Majority of Uterine Leiomyomas. *PLoS One* **7**: e33251.

McNerney ME, Brown CD, Wang X, Bartom ET, Karmakar S, Bandlamudi C, Yu S, Ko J, Sandall BP, Stricker T, Anastasi J, Grossman RL, Cunningham JM, Le Beau MM, White KP (2013) CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia. *Blood* **121**: 975-983.

Mehine M, Heinonen HR, Sarvilinna N, Pitkänen E, Mäkinen N, Katainen R, Tuupanen S, Bützow R, Sjöberg J, Aaltonen LA (2015) Clonally related uterine leiomyomas are common and display branched tumor evolution. *Hum Mol Genet* **24**: 4407-4416.

Mehine M, Kaasinen E, Heinonen HR, Mäkinen N, Kämpjärvi K, Sarvilinna N, Aavikko M, Vähärautio A, Pasanen A, Bützow R, Heikinheimo O, Sjöberg J, Pitkänen E, Vahteristo P, Aaltonen LA (2016) Integrated data analysis reveals uterine leiomyoma subtypes with distinct driver pathways and biomarkers. *Proc Natl Acad Sci U S A* **113**: 1315-1320.

Mehine M, Kaasinen E, Mäkinen N, Katainen R, Kämpjärvi K, Pitkänen E, Heinonen HR, Bützow R, Kilpivaara O, Kuosmanen A, Ristolainen H, Gentile M, Sjöberg J, Vahteristo P, Aaltonen LA (2013) Characterization of Uterine Leiomyomas by Whole-Genome Sequencing. *N Engl J Med* **369**: 43-53.

Mehine M, Mäkinen N, Heinonen HR, Aaltonen LA, Vahteristo P (2014) Genomics of uterine leiomyomas: insights from high-throughput sequencing. *Fertil Steril* **102**: 621-629.

Meloni AM, Surti U, Contento AM, Davare J, Sandberg AA (1992) Uterine leiomyomas: cytogenetic and histologic profile. *Obstet Gynecol* **80**: 209-217.

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* **16**: 1215.

Mishima C, Kagara N, Tanei T, Naoi Y, Shimoda M, Shimomura A, Shimazu K, Kim SJ, Noguchi S (2015) Mutational analysis of MED12 in fibroadenomas and phyllodes tumors of the breast by means of targeted next-generation sequencing. *Breast Cancer Res Treat* **152**: 305-312.

Mittal K and Joutovsky A (2007) Areas with benign morphologic and immunohistochemical features are associated with some uterine leiomyosarcomas. *Gynecol Oncol* **104**: 362-365.

Mittal KR, Chen F, Wei JJ, Rijhvani K, Kurvathi R, Streck D, Dermody J, Toruner GA (2009) Molecular and immunohistochemical evidence for the origin of uterine leiomyosarcomas from associated leiomyoma and symplastic leiomyoma-like areas. *Mod Pathol* **22**: 1303-1311.

Mittal P, Shin YH, Yatsenko SA, Castro CA, Surti U, Rajkovic A (2015) Med12 gain-of-function mutation causes leiomyomas and genomic instability. *J Clin Invest* **125**: 3280-3284.

Moravek MB, Yin P, Ono M, Coon JS 5, Dyson MT, Navarro A, Marsh EE, Chakravarti D, Kim JJ, Wei JJ, Bulun SE (2015) Ovarian steroids, stem cells and uterine leiomyoma: therapeutic implications. *Hum Reprod Update* **21**: 1-12.

Myers LC and Kornberg RD (2000) Mediator of transcriptional regulation. *Annu Rev Biochem* **69**: 729-749.

Mäkinen N, Aavikko M, Heikkinen T, Taipale M, Taipale J, Koivisto-Korander R, Bützow R, Vahteristo P (2016) Exome Sequencing of Uterine Leiomyosarcomas Identifies Frequent Mutations in TP53, ATRX, and MED12. *PLoS Genet* **12**: e1005850.

Mäkinen N, Heinonen HR, Moore S, Tomlinson IP, van der Spuy ZM, Aaltonen LA (2011a) MED12 exon 2 mutations are common in uterine leiomyomas from South African patients. *Oncotarget* **2**: 966-969.

Mäkinen N, Heinonen HR, Sjöberg J, Taipale J, Vahteristo P, Aaltonen LA (2014a) Mutation analysis of components of the Mediator kinase module in MED12 mutation-negative uterine leiomyomas. *Br J Cancer* **110**: 2246-2249.

Mäkinen N, Mehine M, Tolvanen J, Kaasinen E, Li Y, Lehtonen HJ, Gentile M, Yan J, Enge M, Taipale M, Aavikko M, Katainen R, Virolainen E, Böhling T, Koski TA, Launonen V, Sjöberg J, Taipale J, Vahteristo P, Aaltonen LA (2011b) MED12, the mediator complex subunit 12 gene, is mutated at high frequency in uterine leiomyomas. *Science* **334**: 252-255.

Mäkinen N, Vahteristo P, Bützow R, Sjöberg J, Aaltonen LA (2014b) Exomic landscape of MED12 mutation-negative and -positive uterine leiomyomas. *Int J Cancer* **134**: 1008-1012.

Mäkinen N, Vahteristo P, Kämpjärvi K, Arola J, Bützow R, Aaltonen LA (2013) MED12 exon 2 mutations in histopathological uterine leiomyoma variants. *Eur J Hum Genet*: 1300-3.

Nagai R, Brock JW, Blatnik M, Baatz JE, Bethard J, Walla MD, Thorpe SR, Baynes JW, Frizzell N (2007) Succination of protein thiols during adipocyte maturation: a biomarker of mitochondrial stress. *J Biol Chem* **282**: 34219-34228.

Nagasawa S, Maeda I, Fukuda T, Wu W, Hayami R, Kojima Y, Tsugawa K, Ohta T (2015) MED12 exon 2 mutations in phyllodes tumors of the breast. *Cancer Med* **4**: 1117-1121.

Nagy R, Sweet K, Eng C (2004) Highly penetrant hereditary cancer syndromes. *Oncogene* **23**: 6445-6470.

Nakai K and Horton P (1999) PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem Sci* **24**: 34-36.

Negrini S, Gorgoulis VG, Halazonetis TD (2010) Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* **11**: 220-228.

Nelson ND and Bertuch AA (2012) Dyskeratosis congenita as a disorder of telomere maintenance. *Mutat Res* **730**: 43-51.

Neu-Yilik G, Amthor B, Gehring NH, Bahri S, Paidassi H, Hentze MW, Kulozik AE (2011) Mechanism of escape from nonsense-mediated mRNA decay of human beta-globin transcripts with nonsense mutations in the first exon. *RNA* **17**: 843-854.

Ng CC, Tan J, Ong CK, Lim WK, Rajasegaran V, Nasir ND, Lim JC, Thike AA, Salahuddin SA, Iqbal J, Busmanis I, Chong AP, Teh BT, Tan PH (2015) MED12 is frequently mutated in breast phyllodes tumours: a study of 112 cases. *J Clin Pathol* **68**: 685-691.

Nguyen Ba AN, Pogoutse A, Provart N, Moses AM (2009) NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics* **10**: 202-2105-10-202.

Nibert M and Heim S (1990) Uterine leiomyoma cytogenetics. *Genes Chromosomes Cancer* **2**: 3-13.

Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, McBride DJ, Bignell GR, Cooke SL, Shlien A, Gamble J, Whitmore I, Maddison M, Tarpey PS, Davies HR, Papaemmanuil E, Stephens PJ, McLaren S, Butler AP, Teague JW, Jonsson G, Garber JE, Silver D, Miron P, Fatima A, Boyault S, Langerod A, Tutt A, Martens JW, Aparicio SA, Borg A, Salomon AV, Thomas G, Borresen-Dale AL, Richardson AL, Neuberger MS, Futreal PA, Campbell PJ, Stratton MR, Breast Cancer Working Group of the International Cancer Genome Consortium (2012) Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**: 979-993.

Nilbert M, Heim S, Mandahl N, Floderus UM, Willen H, Mitelman F (1990) Characteristic chromosome abnormalities, including rearrangements of 6p, del(7q), +12, and t(12;14), in 44 uterine leiomyomas. *Hum Genet* **85**: 605-611.

Noguchi S, Yokouchi H, Aihara T, Motomura K, Inaji H, Imaoka S, Koyama H (1995) Progression of fibroadenoma to phyllodes tumor demonstrated by clonal analysis. *Cancer* **76**: 1779-1785.

Novellasdemunt L, Antas P, Li VS (2015) Targeting Wnt signaling in colorectal cancer. A Review in the Theme: Cell Signaling: Proteins, Pathways and Mechanisms. *Am J Physiol Cell Physiol* **309**: C511-21.

Näär AM, Taatjes DJ, Zhai W, Nogales E, Tjian R (2002) Human CRSP interacts with RNA polymerase II CTD and adopts a specific CTD-bound conformation. *Genes Dev* **16**: 1339-1344.

Oliva E, Carcangiu ML, Carinelli SG, Ip P, Loening T, Longacre TA, Nucci MR, Prat J, Zaloudek CJ (2014) Mesenchymal tumours. In WHO Classification of Tumours of Female Reproductive Organs, Kurman, R.J., Carcangiu, M.L. et al. (ed) pp 135-147. IARC Press: Lyon.

Ono M, Bulun SE, Maruyama T (2014a) Tissue-specific stem cells in the myometrium and tumor-initiating cells in leiomyoma. *Biol Reprod* **91**: 149.

Ono M, Maruyama T, Masuda H, Kajitani T, Nagashima T, Arase T, Ito M, Ohta K, Uchida H, Asada H, Yoshimura Y, Okano H, Matsuzaki Y (2007) Side population in human uterine myometrium

displays phenotypic and functional characteristics of myometrial stem cells. *Proc Natl Acad Sci U S A* **104**: 18700-18705.

Ono M, Yin P, Navarro A, Moravek MB, Coon JS 5, Druschitz SA, Serna VA, Qiang W, Brooks DC, Malpani SS, Ma J, Ercan CM, Mittal N, Monsivais D, Dyson MT, Yemelyanov A, Maruyama T, Chakravarti D, Kim JJ, Kurita T, Gottardi CJ, Bulun SE (2013) Paracrine activation of WNT/beta-catenin pathway in uterine leiomyoma stem cells promotes tumor growth. *Proc Natl Acad Sci U S A* **110**: 17053-17058.

Ono M, Yin P, Navarro A, Moravek MB, Coon VJS, Druschitz SA, Gottardi CJ, Bulun SE (2014b) Inhibition of canonical WNT signaling attenuates human leiomyoma cell growth. *Fertil Steril* **101**: 1441-1449.

Ooi A, Wong JC, Petillo D, Roossien D, Perrier-Trudova V, Whitten D, Min BW, Tan MH, Zhang Z, Yang XJ, Zhou M, Gardie B, Molinie V, Richard S, Tan PH, Teh BT, Furge KA (2011) An antioxidant response phenotype shared between hereditary and sporadic type 2 papillary renal cell carcinoma. *Cancer Cell* **20**: 511-523.

Opitz JM and Kaveggia EG (1974) Studies of malformation syndromes of man 33: the FG syndrome. An X-linked recessive syndrome of multiple congenital anomalies and mental retardation. *Z Kinderheilkd* **117**: 1-18.

Osinovskaya NS, Malysheva OV, Shved NY, Ivashchenko TE, Sultanov IY, Efimova OA, Yarmolinskaya MI, Bezhenar VF, Baranov VS (2015) Frequency and Spectrum of MED12 Exon 2 Mutations in Multiple Versus Solitary Uterine Leiomyomas From Russian Patients. *Int J Gynecol Pathol* Epub ahead of print.

Packenham JP, du Manoir S, Schrock E, Risinger JI, Dixon D, Denz DN, Evans JA, Berchuck A, Barrett JC, Devereux TR, Ried T (1997) Analysis of genetic alterations in uterine leiomyomas and leiomyosarcomas by comparative genomic hybridization. *Mol Carcinog* **19**: 273-279.

Pallante P, Sepe R, Puca F, Fusco A (2015) High mobility group a proteins as tumor markers. *Front Med (Lausanne)* **2**: 15.

Palles C, Cazier JB, Howarth KM, Domingo E, Jones AM, Broderick P, Kemp Z, Spain SL, Guarino E, Salguero I, Sherborne A, Chubb D, Carvajal-Carmona LG, Ma Y, Kaur K, Dobbins S, Barclay E, Gorman M, Martin L, Kovac MB, Humphray S, CORGI Consortium, WGS500 Consortium, Lucassen A, Holmes CC, Bentley D, Donnelly P, Taylor J, Petridis C, Roylance R, Sawyer EJ, Kerr DJ, Clark S, Grimes J, Kearsey SE, Thomas HJ, McVean G, Houlston RS, Tomlinson I (2013) Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet* **45**: 136-144.

Parazzini F, La Vecchia C, Negri E, Cecchetti G, Fedele L (1988) Epidemiologic characteristics of women with uterine fibroids: a case-control study. *Obstet Gynecol* **72**: 853-857.

Parazzini F, Negri E, La Vecchia C, Chatenoud L, Ricci E, Guarnerio P (1996) Reproductive factors and risk of uterine fibroids. *Epidemiology* **7**: 440-442.

Payne SR and Kemp CJ (2005) Tumor suppressor genetics. *Carcinogenesis* **26**: 2031-2045.

Peng Y, Laser J, Shi G, Mittal K, Melamed J, Lee P, Wei JJ (2008) Antiproliferative effects by Let-7 repression of high-mobility group A2 in uterine leiomyoma. *Mol Cancer Res* **6**: 663-673.

Perot G, Croce S, Ribeiro A, Lagarde P, Velasco V, Neuville A, Coindre JM, Stoeckle E, Floquet A, Macgrogan G, Chibon F (2012) MED12 Alterations in Both Human Benign and Malignant Uterine Soft Tissue Tumors. *PLoS One* **7**: e40015.

Pfarr N, Kriegsmann M, Sinn P, Klauschen F, Endris V, Herpel E, Muckenhuber A, Jesinghaus M, Klosterhalfen B, Penzel R, Lennerz JK, Weichert W, Stenzinger A (2015) Distribution of MED12 mutations in fibroadenomas and phyllodes tumors of the breast--implications for tumor biology and pathological diagnosis. *Genes Chromosomes Cancer* **54**: 444-452.

Philibert RA (2006) A meta-analysis of the association of the HOPA12bp polymorphism and schizophrenia. *Psychiatr Genet* **16**: 73-76.

Philibert RA, Bohle P, Secrest D, Deaderick J, Sandhu H, Crowe R, Black DW (2007) The association of the HOPA(12bp) polymorphism with schizophrenia in the NIMH Genetics Initiative for Schizophrenia sample. *Am J Med Genet B Neuropsychiatr Genet* **144B**: 743-747.

Philibert RA, King BH, Winfield S, Cook EH, Lee YH, Stubblefield B, Damschroder-Williams P, Dea C, Palotie A, Tengstrom C, Martin BM, Ginns EI (1998) Association of an X-chromosome dodecamer insertional variant allele with mental retardation. *Mol Psychiatry* **3**: 303-309.

Philibert RA, Winfield SL, Damschroder-Williams P, Tengstrom C, Martin BM, Ginns EI (1999) The genomic structure and developmental expression patterns of the human OPA-containing gene (HOPA). *Hum Genet* **105**: 174-178.

Piscuoglio S, Murray M, Fusco N, Marchio C, Loo FL, Martelotto LG, Schultheis AM, Akram M, Weigelt B, Brogi E, Reis-Filho JS (2015) MED12 somatic mutations in fibroadenomas and phyllodes tumours of the breast. *Histopathology* **67**: 719-729.

Polakis P (2012) Wnt signaling in cancer. *Cold Spring Harb Perspect Biol* **4**: 10.1101/cshperspect.a008052.

Pollard PJ, Briere JJ, Alam NA, Barwell J, Barclay E, Wortham NC, Hunt T, Mitchell M, Olpin S, Moat SJ, Hargreaves IP, Heales SJ, Chung YL, Griffiths JR, Dalgleish A, McGrath JA, Gleeson MJ, Hodgson SV, Poulsom R, Rustin P, Tomlinson IP (2005) Accumulation of Krebs cycle intermediates and over-expression of HIF1alpha in tumours which result from germline FH and SDH mutations. *Hum Mol Genet* **14**: 2231-2239.

Pollard PJ, Spencer-Dene B, Shukla D, Howarth K, Nye E, El-Bahrawy M, Deheragoda M, Joannou M, McDonald S, Martin A, Igarashi P, Varsani-Brown S, Rosewell I, Poulsom R, Maxwell P, Stamp GW, Tomlinson IP (2007) Targeted inactivation of fh1 causes proliferative renal cyst development and activation of the hypoxia pathway. *Cancer Cell* **11**: 311-319.

Pon JR and Marra MA (2015) Driver and passenger mutations in cancer. *Annu Rev Pathol* **10**: 25-50.

Pritts EA, Parker WH, Olive DL (2009) Fibroids and infertility: an updated systematic review of the evidence. *Fertil Steril* **91**: 1215-1223.

Prontera P, Ottaviani V, Rogaia D, Isidori I, Mencarelli A, Malerba N, Cocciadiferro D, Rolph P, Stangoni G, Vulto-vanSilfhout A, Merla G (2016) A novel MED12 mutation: Evidence for a fourth phenotype. *Am J Med Genet* **170**: 2377-2383.

Puente XS, Bea S, Valdes-Mas R, Villamor N, Gutierrez-Abril J, Martin-Subero JI, Munar M, Rubio-Perez C, Jares P, Aymerich M, Baumann T, Beekman R, Belver L, Carrio A, Castellano G, Clot G, Colado E, Colomer D, Costa D, Delgado J, Enjuanes A, Estivill X, Ferrando AA, Gelpi JL, Gonzalez B, Gonzalez S, Gonzalez M, Gut M, Hernandez-Rivas JM, Lopez-Guerra M, Martin-Garcia D, Navarro A, Nicolas P, Orozco M, Payer AR, Pinyol M, Pisano DG, Puente DA, Queiros AC, Quesada V, Romeo-Casabona CM, Royo C, Royo R, Rozman M, Russinol N, Salaverria I, Stamatopoulos K, Stunnenberg HG, Tamborero D, Terol MJ, Valencia A, Lopez-Bigas N, Torrents D, Gut I, Lopez-Guillermo A, Lopez-Otin C, Campo E (2015) Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **526**: 519-524.

Puente XS, Pinyol M, Quesada V, Conde L, Ordonez GR, Villamor N, Escaramis G, Jares P, Bea S, Gonzalez-Diaz M, Bassaganyas L, Baumann T, Juan M, Lopez-Guerra M, Colomer D, Tubio JM, Lopez C, Navarro A, Tornador C, Aymerich M, Rozman M, Hernandez JM, Puente DA, Freije JM, Velasco G, Gutierrez-Fernandez A, Costa D, Carrio A, Guijarro S, Enjuanes A, Hernandez L, Yague J, Nicolas P, Romeo-Casabona CM, Himmelbauer H, Castillo E, Dohm JC, de Sanjose S, Piris MA, de Alava E, San Miguel J, Royo R, Gelpi JL, Torrents D, Orozco M, Pisano DG, Valencia A, Guigo R, Bayes M, Heath S, Gut M, Klatt P, Marshall J, Raine K, Stebbings LA, Futreal PA, Stratton MR, Campbell PJ, Gut I, Lopez-Guillermo A, Estivill X, Montserrat E, Lopez-Otin C, Campo E (2011) Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **475**: 101-105.

Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D (2011) RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer* **11**: 761-774.

Quade BJ, Weremowicz S, Neskey DM, Vanni R, Ladd C, Dal Cin P, Morton CC (2003) Fusion transcripts involving HMGA2 are not a common molecular mechanism in uterine leiomyomata with rearrangements in 12q15. *Cancer Res* **63**: 1351-1358.

Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, Ramsay AJ, Bea S, Pinyol M, Martinez-Trillos A, Lopez-Guerra M, Colomer D, Navarro A, Baumann T, Aymerich M, Rozman M, Delgado J, Gine E, Hernandez JM, Gonzalez-Diaz M, Puente DA, Velasco G, Freije JM, Tubio JM, Royo R, Gelpi JL, Orozco M, Pisano DG, Zamora J, Vazquez M, Valencia A, Himmelbauer

H, Bayes M, Heath S, Gut M, Gut I, Estivill X, Lopez-Guillermo A, Puente XS, Campo E, Lopez-Otin C (2011) Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet* **44**: 47-52.

Rahman N (2014) Realizing the promise of cancer predisposition genes. *Nature* **505**: 302-308.

Rahman N and Scott RH (2007) Cancer genes associated with phenotypes in monoallelic and biallelic mutation carriers: new lessons from old players. *Hum Mol Genet* **16 Spec No 1**: R60-6.

Rantanen V, Valori M, Hautaniemi S (2014) Anima: modular workflow system for comprehensive image data analysis. *Front Bioeng Biotechnol* **2**: 25.

Rassenti LZ, Huynh L, Toy TL, Chen L, Keating MJ, Gribben JG, Neuberg DS, Flinn IW, Rai KR, Byrd JC, Kay NE, Greaves A, Weiss A, Kipps TJ (2004) ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. *N Engl J Med* **351**: 893-901.

Rassenti LZ, Jain S, Keating MJ, Wierda WG, Grever MR, Byrd JC, Kay NE, Brown JR, Gribben JG, Neuberg DS, He F, Greaves AW, Rai KR, Kipps TJ (2008) Relative value of ZAP-70, CD38, and immunoglobulin mutation status in predicting aggressive disease in chronic lymphocytic leukemia. *Blood* **112**: 1923-1930.

Rau MJ, Fischer S, Neumann CJ (2006) Zebrafish Trap230/Med12 is required as a coactivator for Sox9-dependent neural crest, cartilage and ear development. *Dev Biol* **296**: 83-93.

Ravegnini G, Marino-Enriquez A, Slater J, Eilers G, Wang Y, Zhu M, Nucci MR, George S, Angelini S, Raut CP, Fletcher JA (2013) MED12 mutations in leiomyosarcoma and extrauterine leiomyoma. *Mod Pathol* **26**: 743-749.

Rein MS, Friedman AJ, Barbieri RL, Pavelka K, Fletcher JA, Morton CC (1991) Cytogenetic abnormalities in uterine leiomyomata. *Obstet Gynecol* **77**: 923-926.

Rieker RJ, Agaimy A, Moskalev EA, Hebele S, Hein A, Mehlhorn G, Beckmann MW, Hartmann A, Haller F (2013) Mutation status of the mediator complex subunit 12 (MED12) in uterine leiomyomas and concurrent/metachronous multifocal peritoneal smooth muscle nodules (leiomyomatosis peritonealis disseminata). *Pathology* **45**: 388-392.

Risheg H, Graham JM,Jr, Clark RD, Rogers RC, Opitz JM, Moeschler JB, Peiffer AP, May M, Joseph SM, Jones JR, Stevenson RE, Schwartz CE, Friez MJ (2007) A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat Genet* **39**: 451-453.

Robinson D, Van Allen EM, Wu YM, Schultz N, Lonigro RJ, Mosquera JM, Montgomery B, Taplin ME, Pritchard CC, Attard G, Beltran H, Abida W, Bradley RK, Vinson J, Cao X, Vats P, Kunju LP, Hussain M, Feng FY, Tomlins SA, Cooney KA, Smith DC, Brennan C, Siddiqui J, Mehra R, Chen Y, Rathkopf DE, Morris MJ, Solomon SB, Durack JC, Reuter VE, Gopalan A, Gao J, Loda M, Lis RT, Bowden M, Balk SP, Gaviola G, Sougnez C, Gupta M, Yu EY, Mostaghel EA, Cheng HH, Mulcahy H, True LD, Plymate SR, Dvinge H, Ferraldeschi R, Flohr P, Miranda S, Zafeiriou Z, Tunariu N, Mateo J, Perez-Lopez R, Demichelis F, Robinson BD, Schiffman M, Nanus DM, Tagawa ST, Sigaras A, Eng KW, Elemento O, Sboner A, Heath EI, Scher HI, Pienta KJ, Kantoff P, de Bono JS, Rubin MA, Nelson PS, Garraway LA, Sawyers CL, Chinnaiyan AM (2015) Integrative clinical genomics of advanced prostate cancer. *Cell* **161**: 1215-1228.

Rocha PP, Scholze M, Bleiss W, Schrewe H (2010) Med12 is essential for early mouse development and for canonical Wnt and Wnt/PCP signaling. *Development* **137**: 2723-2731.

Ross RK, Pike MC, Vessey MP, Bull D, Yeates D, Casagrande JT (1986) Risk factors for uterine fibroids: reduced risk associated with oral contraceptives. *Br Med J (Clin Res Ed)* **293**: 359-362.

Rossi D and Gaidano G (2016) The clinical implications of gene mutations in chronic lymphocytic leukaemia. *Br J Cancer* **114**: 849-854.

Roux KJ, Kim DI, Raida M, Burke B (2012) A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* **196**: 801-810.

Rump P, Niessen RC, Verbruggen KT, Brouwer OF, de Raad M, Hordijk R (2011) A novel mutation in MED12 causes FG syndrome (Opitz-Kaveggia syndrome). *Clin Genet* **79**: 183-188.

Sadeghi S, Khorrami M, Amin-Beidokhti M, Abbasi M, Kamalian Z, Irani S, Omrani M, Azmoodeh O, Mirfakhraie R (2016) The study of MED12 gene mutations in uterine leiomyomas from Iranian patients. *Tumour Biol* **37**: 1567-1571.

Salovaara R, Loukola A, Kristo P, Kaariainen H, Ahtola H, Eskelinen M, Harkonen N, Julkunen R, Kangas E, Ojala S, Tulikoura J, Valkamo E, Järvinen H, Mecklin JP, Aaltonen LA, de la Chapelle A (2000) Population-based molecular detection of hereditary nonpolyposis colorectal cancer. *J Clin Oncol* **18**: 2193-2200.

Sandberg AA (2005a) Updates on the cytogenetics and molecular genetics of bone and soft tissue tumors: leiomyoma. *Cancer Genet Cytogenet* **158**: 1-26.

Sandberg AA (2005b) Updates on the cytogenetics and molecular genetics of bone and soft tissue tumors: leiomyosarcoma. *Cancer Genet Cytogenet* **161**: 1-19.

Sanz-Ortega J, Vocke C, Stratton P, Linehan WM, Merino MJ (2013) Morphologic and molecular characteristics of uterine leiomyomas in hereditary leiomyomatosis and renal cancer (HLRCC) syndrome. *Am J Surg Pathol* **37**: 74-80.

Sato F, Miyake H, Nishi M, Mori M, Kudo R (2000) Early normal menstrual cycle pattern and the development of uterine leiomyomas. *J Womens Health Gend Based Med* **9**: 299-302.

Sato F, Mori M, Nishi M, Kudo R, Miyake H (2002) Familial aggregation of uterine myomas in Japanese women. *J Epidemiol* **12**: 249-253.

Sato S, Tomomori-Sato C, Parmely TJ, Florens L, Zybailov B, Swanson SK, Banks CA, Jin J, Cai Y, Washburn MP, Conaway JW, Conaway RC (2004) A set of consensus mammalian mediator subunits identified by multidimensional protein identification technology. *Mol Cell* **14**: 685-691.

Schoenmakers EF, Bunt J, Hermers L, Schepens M, Merkx G, Janssen B, Kersten M, Huys E, Pauwels P, Debiec-Rychter M, van Kessel AG (2013) Identification of CUX1 as the recurrent chromosomal band 7q22 target gene in human uterine leiomyoma. *Genes Chromosomes Cancer* **52**: 11-23.

Schoenmakers EF, Huysmans C, Van de Ven WJ (1999) Allelic knockout of novel splice variants of human recombination repair gene RAD51B in t(12;14) uterine leiomyomas. *Cancer Res* **59**: 19-23.

Schoenmakers EF, Mols R, Wanschura S, Kools PF, Geurts JM, Bartnitzke S, Bullerdiek J, van den Berghe H, Van de Ven WJ (1994) Identification, molecular cloning, and characterization of the chromosome 12 breakpoint cluster region of uterine leiomyomas. *Genes Chromosomes Cancer* **11**: 106-118.

Schoenmakers EF, Wanschura S, Mols R, Bullerdiek J, Van den Berghe H, Van de Ven WJ (1995) Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours. *Nat Genet* **10**: 436-444.

Schuh A, Becq J, Humphray S, Alexa A, Burns A, Clifford R, Feller SM, Grocock R, Henderson S, Khrebtukova I, Kingsbury Z, Luo S, McBride D, Murray L, Menju T, Timbs A, Ross M, Taylor J, Bentley D (2012) Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood* **120**: 4191-4196.

Schuster-Bockler B and Lehner B (2012) Chromatin organization is a major influence on regional mutation rates in human cancer cells. *Nature* **488**: 504-507.

Schwartz CE, Tarpey PS, Lubs HA, Verloes A, May MM, Risheg H, Friez MJ, Futreal PA, Edkins S, Teague J, Briault S, Skinner C, Bauer-Carlin A, Simensen RJ, Joseph SM, Jones JR, Gecz J, Stratton MR, Raymond FL, Stevenson RE (2007) The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. *J Med Genet* **44**: 472-477.

Schwetye KE, Pfeifer JD, Duncavage EJ (2014) MED12 exon 2 mutations in uterine and extrauterine smooth muscle tumors. *Hum Pathol* **45**: 65-70.

Sehgal R, Sheahan K, O'Connell PR, Hanly AM, Martin ST, Winter DC (2014) Lynch syndrome: an updated review. *Genes (Basel)* **5**: 497-507.

Shahbazi S, Fatahi N, Amini-Moghaddam S (2015) Somatic mutational analysis of MED12 exon 2 in uterine leiomyomas of Iranian women. *Am J Cancer Res* **5**: 2441-2446.

Shaikhibrahim Z, Offermann A, Braun M, Menon R, Syring I, Nowak M, Halbach R, Vogel W, Ruiz C, Zellweger T, Rentsch CA, Svensson M, Andren O, Bubendorf L, Biskup S, Duensing S, Kirfel J, Perner S (2014) MED12 overexpression is a frequent event in castration-resistant prostate cancer. *Endocr Relat Cancer* **21**: 663-675.

Shikora SA, Niloff JM, Bistrian BR, Forse RA, Blackburn GL (1991) Relationship between obesity and uterine leiomyomata. *Nutrition* **7**: 251-255.

Shin CH, Chung WS, Hong SK, Ober EA, Verkade H, Field HA, Huisken J, Stainier DY (2008) Multiple roles for Med12 in vertebrate endoderm development. *Dev Biol* **317**: 467-479.

Smit DL, Mensenkamp AR, Badeloe S, Breuning MH, Simon ME, van Spaendonck KY, Aalfs CM, Post JG, Shanley S, Krapels IP, Hoefsloot LH, van Moorselaar RJ, Starink TM, Bayley JP, Frank J, van Steensel MA, Menko FH (2011) Hereditary leiomyomatosis and renal cell cancer in families referred for fumarate hydratase germline mutation analysis. *Clin Genet* **79**: 49-59.

Spinella J, Cassart P, Richer C, Saillour V, Ouimet M, Langlois S, St-Onge P, Sontag T, Healy J, Sinnett D (2016) The genomic landscape of pediatric T-cell acute lymphoblastic leukemia reveals novel X-linked somatic mutations associated with apoptosis resistance. *Oncotarget* Epub ahead of print.

Staats B, Bonk U, Wanschura S, Hanisch P, Schoenmakers EF, Van de Ven WJ, Bartnitzke S, Bullerdiek J (1996) A fibroadenoma with a t(4;12) (q27;q15) affecting the HMGI-C gene, a member of the high mobility group protein gene family. *Breast Cancer Res Treat* **38**: 299-303.

Stamatoyannopoulos JA, Adzhubei I, Thurman RE, Kryukov GV, Mirkin SM, Sunyaev SR (2009) Human mutation rate associated with DNA replication timing. *Nat Genet* **41**: 393-395.

Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, McLaren S, Lin ML, McBride DJ, Varela I, Nik-Zainal S, Leroy C, Jia M, Menzies A, Butler AP, Teague JW, Quail MA, Burton J, Swerdlow H, Carter NP, Morsberger LA, Iacobuzio-Donahue C, Follows GA, Green AR, Flanagan AM, Stratton MR, Futreal PA, Campbell PJ (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**: 27-40.

Stewart EA (2015) Clinical practice. Uterine fibroids. *N Engl J Med* **372**: 1646-1655.

Stewart EA, Laughlin-Tommaso SK, Catherino WH, Lalitkumar S, Gupta D, Vollenhoven B (2016) Uterine fibroids. *Nat Rev Dis Primers* **2**: 16043.

Stewart L, Glenn GM, Stratton P, Goldstein AM, Merino MJ, Tucker MA, Linehan WM, Toro JR (2008) Association of germline mutations in the fumarate hydratase gene and uterine fibroids in women with hereditary leiomyomatosis and renal cell cancer. *Arch Dermatol* **144**: 1584-1592.

Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, Kryukov GV, Lawrence MS, Sougnez C, McKenna A, Shefler E, Ramos AH, Stojanov P, Carter SL, Voet D, Cortes ML, Auclair D, Berger MF, Saksena G, Guiducci C, Onofrio RC, Parkin M, Romkes M, Weissfeld JL, Seethala RR, Wang L, Rangel-Escareno C, Fernandez-Lopez JC, Hidalgo-Miranda A, Melendez-Zajgla J, Winckler W, Ardlie K, Gabriel SB, Meyerson M, Lander ES, Getz G, Golub TR, Garraway LA, Grandis JR (2011) The mutational landscape of head and neck squamous cell carcinoma. *Science* **333**: 1157-1160.

Summers WE, Watson RL, Wooldridge WH, Langford HG (1971) Hypertension, obesity, and fibromyomata uteri, as a syndrome. *Arch Intern Med* **128**: 750-754.

Suwaki N, Klare K, Tarsounas M (2011) RAD51 paralogs: roles in DNA damage signalling, recombinational repair and tumorigenesis. *Semin Cell Dev Biol* **22**: 898-905.

Sveen A, Kilpinen S, Ruusulehto A, Lothe RA, Skotheim RI (2016) Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene* **35**: 2413-2427.

Swanton C (2012) Intratumor heterogeneity: evolution through space and time. *Cancer Res* **72**: 4875-4882.

Taguchi K, Motohashi H, Yamamoto M (2011) Molecular mechanisms of the Keap1-Nrf2 pathway in stress response and cancer evolution. *Genes Cells* **16**: 123-140.

Tallini G, Vanni R, Manfioletti G, Kazmierczak B, Faa G, Pauwels P, Bullerdiek J, Giancotti V, Van Den Berghe H, Dal Cin P (2000) HMGI-C and HMGI(Y) immunoreactivity correlates with cytogenetic abnormalities in lipomas, pulmonary chondroid hamartomas, endometrial polyps, and uterine leiomyomas and is compatible with rearrangement of the HMGI-C and HMGI(Y) genes. *Lab Invest* **80**: 359-369.

Tan PH, Thike AA, Tan WJ, Thu MM, Busmanis I, Li H, Chay WY, Tan MH, Phyllodes Tumour Network Singapore (2012) Predicting clinical behaviour of breast phyllodes tumours: a nomogram based on histological criteria and surgical margins. *J Clin Pathol* **65**: 69-76.

Tan WJ, Chan JY, Thike AA, Lim JC, Md Nasir ND, Tan JS, Koh VC, Lim WK, Tan J, Ng CC, Rajasegaran V, Nagarajan S, Bay BH, Teh BT, Tan PH (2016) MED12 protein expression in breast fibroepithelial lesions: correlation with mutation status and oestrogen receptor expression. *J Clin Pathol* Epub ahead of print.

Tanwar PS, Lee HJ, Zhang L, Zukerberg LR, Taketo MM, Rueda BR, Teixeira JM (2009) Constitutive activation of Beta-catenin in uterine stroma and smooth muscle leads to the development of mesenchymal tumors in mice. *Biol Reprod* **81**: 545-552.

Tavassoli FA and Devilee P (eds) (2003) World Health Organization Classification of Tumours. Pathology and Genetics of the Breast and Female Genital Organs. IARC Press: Lyon.

Terry KL, De Vivo I, Hankinson SE, Missmer SA (2010) Reproductive characteristics and risk of uterine leiomyomata. *Fertil Steril* **94**: 2703-2707.

Thiagalingam S, Lengauer C, Leach FS, Schutte M, Hahn SA, Overhauser J, Willson JK, Markowitz S, Hamilton SR, Kern SE, Kinzler KW, Vogelstein B (1996) Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. *Nat Genet* **13**: 343-346.

Thielen BK, Barker DF, Nelson RD, Zhou J, Kren SM, Segal Y (2003) Deletion mapping in Alport syndrome and Alport syndrome-diffuse leiomyomatosis reveals potential mechanisms of visceral smooth muscle overgrowth. *Hum Mutat* **22**: 419.

Tolvanen J, Uimari O, Ryynanen M, Aaltonen LA, Vahteristo P (2012) Strong family history of uterine leiomyomatosis warrants fumarate hydratase mutation screening. *Hum Reprod* **27**: 1865-1869.

Tomlinson IP, Alam NA, Rowan AJ, Barclay E, Jaeger EE, Kelsell D, Leigh I, Gorman P, Lamlum H, Rahman S, Roylance RR, Olpin S, Bevan S, Barker K, Hearle N, Houlston RS, Kiuru M, Lehtonen R, Karhu A, Vilkki S, Laiho P, Eklund C, Vierimaa O, Aittomäki K, Hietala M, Sistonen P, Paetau A, Salovaara R, Herva R, Launonen V, Aaltonen LA, Multiple Leiomyoma Consortium (2002) Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nat Genet* **30**: 406-410.

Toro JR, Nickerson ML, Wei MH, Warren MB, Glenn GM, Turner ML, Stewart L, Duray P, Tourre O, Sharma N, Choyke P, Stratton P, Merino M, Walther MM, Linehan WM, Schmidt LS, Zbar B (2003) Mutations in the fumarate hydratase gene cause hereditary leiomyomatosis and renal cell cancer in families in North America. *Am J Hum Genet* **73**: 95-106.

Toro JR, Travis LB, Wu HJ, Zhu K, Fletcher CD, Devesa SS (2006) Incidence patterns of soft tissue sarcomas, regardless of primary site, in the surveillance, epidemiology and end results program, 1978-2001: An analysis of 26,758 cases. *Int J Cancer* **119**: 2922-2930.

Townsend DE, Sparkes RS, Baluda MC, McClelland G (1970) Unicellular histogenesis of uterine leiomyomas as determined by electrophoresis by glucose-6-phosphate dehydrogenase. *Am J Obstet Gynecol* **107**: 1168-1173.

Tsai KL, Sato S, Tomomori-Sato C, Conaway RC, Conaway JW, Asturias FJ (2013) A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. *Nat Struct Mol Biol* **20**: 611-619.

Tsai KL, Tomomori-Sato C, Sato S, Conaway RC, Conaway JW, Asturias FJ (2014) Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. *Cell* **157**: 1430-1444.

Tse GM, Lee CS, Kung FY, Scolyer RA, Law BK, Lau TS, Putti TC (2002) Hormonal receptors expression in epithelial cells of mammary phyllodes tumors correlates with pathologic grade of the tumor: a multicenter study of 143 cases. *Am J Clin Pathol* **118**: 522-526.

Tsutsui T, Fukasawa R, Tanaka A, Hirose Y, Okhuma Y (2011) Identification of target genes for the CDK subunits of the Mediator complex. *Genes Cells* **16**: 1208-1218.

Tsutsui T, Umemura H, Tanaka A, Mizuki F, Hirose Y, Ohkuma Y (2008) Human mediator kinase subunit CDK11 plays a negative role in viral activator VP16-dependent transcriptional regulation. *Genes Cells* **13**: 817-826.

Turunen M, Spaeth JM, Keskitalo S, Park MJ, Kivioja T, Clark AD, Mäkinen N, Gao F, Palin K, Nurkkala H, Vähärautio A, Aavikko M, Kämpjärvi K, Vahteristo P, Kim CA, Aaltonen LA, Varjosalo M, Taipale J, Boyer TG (2014) Uterine Leiomyoma-Linked MED12 Mutations Disrupt Mediator-Associated CDK Activity. *Cell Rep* **7**: 654-660.

UniProt Consortium (2015) UniProt: a hub for protein information. *Nucleic Acids Res* **43**: D204-12.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**: e115.

Vahteristo P, Koski TA, Naatsaari L, Kiuru M, Karhu A, Herva R, Sallinen SL, Vierimaa O, Bjorck E, Richard S, Gardie B, Bessis D, Van Glabeke E, Blanco I, Houlston R, Senter L, Hietala M, Aittomäki K, Aaltonen LA, Launonen V, Lehtonen R (2010) No evidence for a genetic modifier for renal cell cancer risk in HLRCC syndrome. *Fam Cancer* **9**: 245-251.

Van de Ven WJ (1998) Genetic basis of uterine leiomyoma: involvement of high mobility group protein genes. *Eur J Obstet Gynecol Reprod Biol* **81**: 289-293.

Vanharanta S, Pollard PJ, Lehtonen HJ, Laiho P, Sjöberg J, Leminen A, Aittomäki K, Arola J, Kruhoffer M, Orntoft TF, Tomlinson IP, Kiuru M, Arango D, Aaltonen LA (2006) Distinct expression profile in fumarate-hydratase-deficient uterine fibroids. *Hum Mol Genet* **15**: 97-103.

Vanni R, Dal Cin P, Marras S, Moerman P, Andria M, Valdes E, Deprest J, Van den Berghe H (1993) Endometrial polyp: another benign tumor characterized by 12q13-q15 changes. *Cancer Genet Cytogenet* **68**: 32-33.

Vanni R, Lecca U, Faa G (1991) Uterine leiomyoma cytogenetics. II. Report of forty cases. *Cancer Genet Cytogenet* **53**: 247-256.

Vanni R, Nieddu M, Paoli R, Lecca U (1989) Uterine leiomyoma cytogenetics. I. Rearrangements of chromosome 12. *Cancer Genet Cytogenet* **37**: 49-54.

Vanni R, Van Roy N, Lecca U, Speleman F (1992) Uterine leiomyoma cytogenetics. III. Interphase cytogenetic analysis of karyotypically normal uterine leiomyoma excludes possibility of undetected trisomy 12. *Cancer Genet Cytogenet* **62**: 40-42.

Varjosalo M, Keskitalo S, Van Drogen A, Nurkkala H, Vichalkovski A, Aebersold R, Gstaiger M (2013) The protein interaction landscape of the human CMGC kinase group. *Cell Rep* **3**: 1306-1320.

Vikhlyaeva EM, Khodzhaeva ZS, Fantschenko ND (1995) Familial predisposition to uterine leiomyomas. *Int J Gynaecol Obstet* **51**: 127-131.

Vogelstein B and Kinzler KW (2015) The Path to Cancer --Three Strikes and You're Out. *N Engl J Med* **373**: 1895-1898.

Vogelstein B and Kinzler KW (2004) Cancer genes and the pathways they control. *Nat Med* **10**: 789-799.

Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA,Jr, Kinzler KW (2013) Cancer genome landscapes. *Science* **339**: 1546-1558.

Vogl MR, Reiprich S, Kuspert M, Kosian T, Schrewe H, Nave KA, Wegner M (2013) Sox10 cooperates with the mediator subunit 12 during terminal differentiation of myelinating glia. *J Neurosci* **33**: 6679-6690.

Vulto-van Silfhout AT, de Vries BB, van Bon BW, Hoischen A, Ruiterkamp-Versteeg M, Gilissen C, Gao F, van Zwam M, Harteveld CL, van Essen AJ, Hamel BC, Kleefstra T, Willemsen MA, Yntema HG, van Bokhoven H, Brunner HG, Boyer TG, de Brouwer AP (2013) Mutations in MED12 Cause X-Linked Ohdo Syndrome. *Am J Hum Genet* **92**: 401-406.

Wang H, Ye J, Qian H, Zhou R, Jiang J, Ye L (2015a) High-resolution melting analysis of MED12 mutations in uterine leiomyomas in Chinese patients. *Genet Test Mol Biomarkers* **19**: 162-166.

Wang K, Kan J, Yuen ST, Shi ST, Chu KM, Law S, Chan TL, Kan Z, Chan AS, Tsui WY, Lee SP, Ho SL, Chan AK, Cheng GH, Roberts PC, Rejto PA, Gibson NW, Pocalyko DJ, Mao M, Xu J, Leung SY (2011a) Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat Genet* **43**: 1219-1223.

Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, Werner L, Sivachenko A, DeLuca DS, Zhang L, Zhang W, Vartanov AR, Fernandes SM, Goldstein NR, Folco EG, Cibulskis K, Tesar B, Sievers QL, Shefler E, Gabriel S, Hacohen N, Reed R, Meyerson M, Golub TR, Lander ES, Neuberg D, Brown JR, Getz G, Wu CJ (2011b) SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med* **365**: 2497-2506.

Wang L, Shalek AK, Lawrence M, Ding R, Gaublomme JT, Pochet N, Stojanov P, Sougnez C, Shukla SA, Stevenson KE, Zhang W, Wong J, Sievers QL, MacDonald BT, Vartanov AR, Goldstein NR, Neuberg D, He X, Lander E, Hacohen N, Regev A, Getz G, Brown JR, Park H, Wu CJ (2014a) Somatic mutation as a mechanism of Wnt/beta-catenin pathway activation in CLL. *Blood* **124**: 1089-1098.

Wang L, Zeng H, Wang Q, Zhao Z, Boyer TG, Bian X, Xu W (2015b) MED12 methylation by CARM1 sensitizes human breast cancer cells to chemotherapy drugs. *Sci Adv* **1**: e1500463.

Wang Q, Lasset C, Desseigne F, Frappaz D, Bergeron C, Navarro C, Ruano E, Puisieux A (1999) Neurofibromatosis and early onset of cancers in hMLH1-deficient children. *Cancer Res* **59**: 294-297.

Wang X, Sun Q, Ding Z, Ji J, Wang J, Kong X, Yang J, Cai G (2014b) Redefining the modular organization of the core Mediator complex. *Cell Res* **24**: 796-808.

Wang X, Yang N, Uno E, Roeder RG, Guo S (2006) A subunit of the mediator complex regulates vertebrate neuronal development. *Proc Natl Acad Sci U S A* **103**: 17284-17289.

Watson IR, Takahashi K, Futreal PA, Chin L (2013) Emerging patterns of somatic mutations in cancer. *Nat Rev Genet* **14**: 703-718.

Wei MH, Toure O, Glenn GM, Pithukpakorn M, Neckers L, Stolle C, Choyke P, Grubb R, Middelton L, Turner ML, Walther MM, Merino MJ, Zbar B, Linehan WM, Toro JR (2006) Novel mutations in FH and expansion of the spectrum of phenotypes expressed in families with hereditary leiomyomatosis and renal cell cancer. *J Med Genet* **43**: 18-27.

Weinberg RA (1994) Oncogenes and tumor suppressor genes. *CA Cancer J Clin* **44**: 160-170.

Willis A, Jung EJ, Wakefield T, Chen X (2004) Mutant p53 exerts a dominant negative effect by preventing wild-type p53 from binding to the promoter of its target genes. *Oncogene* **23**: 2330-2338.

Wise LA, Palmer JR, Spiegelman D, Harlow BL, Stewart EA, Adams-Campbell LL, Rosenberg L (2005) Influence of body size and body fat distribution on risk of uterine leiomyomata in U.S. black women. *Epidemiology* **16**: 346-354.

Wong CC, Martincorena I, Rust AG, Rashid M, Alifrangis C, Alexandrov LB, Tiffen JC, Kober C, Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium, Green AR, Massie CE, Nangalia J, Lempidaki S, Dohner H, Dohner K, Bray SJ, McDermott U, Papaemmanuil E, Campbell PJ, Adams DJ (2014) Inactivating CUX1 mutations promote tumorigenesis. *Nat Genet* **46**: 33-38.

Yang X, Kandil D, Cosar EF, Khan A (2014) Fibroepithelial tumors of the breast: pathologic and immunohistochemical features and molecular mechanisms. *Arch Pathol Lab Med* **138**: 25-36.

Yin JW and Wang G (2014) The Mediator complex: a master coordinator of transcription and cell lineage development. *Development* **141**: 977-987.

Ylisaukko-oja SK, Kiuru M, Lehtonen HJ, Lehtonen R, Pukkala E, Arola J, Launonen V, Aaltonen LA (2006) Analysis of fumarate hydratase mutations in a population-based series of early onset uterine leiomyosarcoma patients. *Int J Cancer* **119**: 283-287.

Yoon N, Bae GE, Kang SY, Choi MS, Hwang HW, Kim SW, Lee JE, Nam SJ, Gong G, Lee HJ, Bae YK, Lee A, Cho EY (2016) Frequency of MED12 mutations in phyllodes tumors: Inverse correlation with histologic grade. *Genes Chromosomes Cancer* **55**: 495-504.

Yoshida M, Sekine S, Ogawa R, Yoshida H, Maeshima A, Kanai Y, Kinoshita T, Ochiai A (2015) Frequent MED12 mutations in phyllodes tumours of the breast. *Br J Cancer* **112**: 1703-1708.

Zeng WR, Scherer SW, Koutsilieris M, Huizenga JJ, Filteau F, Tsui LC, Nepveu A (1997) Loss of heterozygosity and reduced expression of the CUTL1 gene in uterine leiomyomas. *Oncogene* **14**: 2355-2365.

Zhang L and Long X (2015) Association of BRCA1 promoter methylation with sporadic breast cancers: Evidence from 40 studies. *Sci Rep* **5**: 17869.

Zhang P, Zhang C, Hao J, Sung CJ, Quddus MR, Steinhoff MM, Lawrence WD (2006) Use of X-chromosome inactivation pattern to determine the clonal origins of uterine leiomyoma and leiomyosarcoma. *Hum Pathol* **37**: 1350-1356.

Zhang Q, Ubago J, Li L, Guo H, Liu Y, Qiang W, Kim JJ, Kong B, Wei JJ (2014) Molecular analyses of 6 different types of uterine smooth muscle tumors: Emphasis in atypical leiomyoma. *Cancer* **120**: 3165-3177.

Zhou H, Kim S, Ishii S, Boyer TG (2006) Mediator modulates Gli3-dependent Sonic hedgehog signaling. *Mol Cell Biol* **26**: 8667-8682.

Zhou H, Spaeth JM, Kim NH, Xu X, Friez MJ, Schwartz CE, Boyer TG (2012) MED12 mutations link intellectual disability syndromes with dysregulated GLI3-dependent Sonic Hedgehog signaling. *Proc Natl Acad Sci U S A* **109**: 19763-19768.

Zhou R, Bonneaud N, Yuan CX, de Santa Barbara P, Boizet B, Schomber T, Scherer G, Roeder RG, Poulat F, Berta P (2002) SOX9 interacts with a component of the human thyroid hormone receptor-associated protein complex. *Nucleic Acids Res* **30**: 3245-3252.

Zinn AB, Kerr DS, Hoppel CL (1986) Fumarase deficiency: a new cause of mitochondrial encephalomyopathy. *N Engl J Med* **315**: 469-475.

Websites and public databases:

Finnish Cancer Registry, statistics	www.cancer.fi/syoparekisteri/en/statistics/
Ensemble	www.ensembl.org
UniProt	www.uniprot.org
Pfam	pfam.xfam.org
COSMIC	cancer.sanger.ac.uk/cosmic
CBioPortal	www.cbioportal.org
OMIM	www.omim.org
Medisapiens	medisapiens.com
TCGA	cancergenome.nih.gov
ICGC	icgc.org
Primer3Plus	www.primer3plus.com
FinchTV	www.geospiza.com/ftvdlinfo
PolyPhen2	genetics.bwh.harvard.edu/pph2/
SIFT	sift.jcvi.org/
MUSCLE	www.ebi.ac.uk/Tools/msa/muscle/
SeqNLS	mleg.cse.sc.edu/seqNLS/
PSORT II	psort.hgc.jp/
cNLS Mapper	nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi
NLStradamus	www.moseslab.csb.utoronto.ca/NLStradamus/
UniProtKB/SwissProt	www.uniprot.org/uniprot/
R foundation for statistical computing	www.r-project.org
Python software foundation	www.python.org
NCBI's gene expression archive	www.ncbi.nlm.nih.gov/geo/